Université Paris Cité

Ecole doctorale Bio Sorbonne Paris Cité (ED562)
Laboratoire UMR_S1016 CNRS 8104 Institut Cochin
Equipe Tolérance, biomarqueurs et thérapies dans le diabète de type 1

The antigenic visibility of β cells as a driver of type 1 diabetes: insights from the HLA peptidome

Par Alexia Carré

Thèse de doctorat d'Immunologie

Dirigée par Roberto Mallone

Présentée et soutenue publiquement le 25 novembre 2022

Devant un jury composé de :

Dr. Wei Wu, Associate professor, Université Nationale de Singapour	Rapportrice
Pr. Miriam Cnop, Full professor, Université Libre de Bruxelles	Rapportrice
Pr. Didier Hober, PU-PH, Université de Lille	Examinateur
Pr. David Price, Full professor, Université de Cardiff	Examinateur
Dr. Loredana Saveanu, DR, Université Paris Cité	Examinatrice
Dr. Nicola Ternette, Associate professor, Université d'Oxford	Invitée
Dr. Sylvaine You, DR, Université Paris Cité	Invitée
Pr. Roberto Mallone, PU-PH, Université Paris Cité	Directeur de thèse

Abstract

The antigenic visibility of β cells as a driver of type 1 diabetes: insights from the HLA peptidome

Type 1 diabetes (T1D) is an autoimmune disease, which results from the destruction of insulin-producing β cells by self-reactive $CD8^+$ T cells. These CD8+ T cells recognize class I peptide-HLA (pHLA-I) complexes expressed on the surface of β cells, and our laboratory has demonstrated that their frequency is similar in the blood of diabetic and healthy donors. These results suggest that a universal state of 'benign' autoimmunity exists, and question the mechanisms that drive T1D onset. The switch toward ‘pathological’ autoimmunity could rely on the β cells themselves, which might become more visible and vulnerable to the immune system under some stressful conditions. Moreover, β cells secrete not only insulin, but also extracellular vesicles, exosomes, intact insulin granules, and other proteins contained in the insulin granules. All these components can possibly be phagocytosed, processed, and presented by dendritic cells in the pancreatic lymph nodes, thus triggering the activation of self-reactive T cells.

My project therefore aimed at investigating the visibility of the β cell under different types of external stressors, by characterizing the peptides presented at their surface within the pHLA-I complexes (immunopeptidome) and by assessing the contribution of certain molecular actors to their formation. More specifically, I have studied the effect of Coxsackievirus B (CVB) infection, which is one of the main environmental suspects in T1D, as well as the role of interferon (IFN)-α, which is the earliest signature cytokine observed in children at risk of developing the disease and which is produced locally post-infection.

By infecting a β cell line with different strains of CVB, we made the striking observation that the virus reduces the visibility of the β cell and only few viral peptides are presented. We also demonstrated that the $CD8^+$ T cell response to these sequences is limited. In contrast, exposing β cells to IFN-α greatly increases their visibility. The amount of pHLA complexes presented is higher than in unstimulated cells, peptides restricted to HLA-B alleles are enriched and peptide diversity is increased. Some neo-epitopes were found to result from alternative splicing at the RNA level, at the peptide level, or from post-translational modifications. Of note, the proteasome and immunoproteasome have been identified as key players in the formation of many of these sequences.

Collectively, these results provide a comprehensive mapping of antigenic targets exposed by β cells for recognition by autoimmune T cells. These antigens offer the possibility of tracking autoimmune and anti-CVB responses during the natural history of T1D.

Keywords: Type 1 diabetes, human leukocyte antigen, immunopeptidomics, autoimmunity, Coxsackievirus B, IFN-α

Résumé

La visibilité antigénique de la cellule bêta contribue-t-elle au diabète de type 1 ? Analyse par peptidomique HLA

Le diabète de type 1 (DT1) est une maladie auto-immune, qui résulte de la destruction des cellules bêta, productrices d'insuline, par des lymphocytes T (LT) CD8+ auto-réactifs. Ces LT CD8+ reconnaissent les complexes peptides-HLA de classe I (pHLA-I) exprimés à la surface des cellules bêta et notre laboratoire a montré que leur fréquence est similaire dans le sang de donneurs diabétiques et sains. Ces résultats suggèrent qu'un état universel d'auto-immunité « bénigne » existe, et interroge sur les mécanismes qui gouvernent le déclenchement du DT1. Ce changement d'état pourrait dépendre des cellules bêta elles-mêmes, qui deviendraient davantage visibles au système immunitaire. De plus, plusieurs autres protéines contenues dans les granules d'insuline sont sécrétées avec l'insuline elle-même, sous forme soluble ou dans des vésicules extracellulaires, exosomes, granules entiers. Ces composants peuvent être phagocytés, apprêtés et présentés par des cellules dendritiques dans les ganglions pancréatiques, et ainsi mener à l'activation des LT auto-réactifs.

Mes travaux de thèse ont donc porté sur l'étude de la visibilité de la cellule bêta soumise à différents stress, en caractérisant les peptides présentés à sa surface au sein des complexes pHLA-I (immunopeptidome) et en évaluant la contribution de certains acteurs moléculaires à leur formation. Plus précisément, je me suis intéressée à l'un des facteurs environnementaux suspectés de jouer un rôle dans le développement de la maladie, le Coxsackievirus B (CVB), ainsi qu'à l'effet de l'interféron-alpha, une cytokine signature observée précocement chez les individus à risque de développer le DT1 et qui est produite localement après infection.

En infectant une lignée cellulaire bêta avec différentes souches de CVB, nous avons observé que le virus réduit la visibilité de la cellule bêta et peu de peptides viraux sont présentés. Nous avons aussi démontré que la réponse LT CD8+ à ces séquences est limitée et majoritairement naïve. A l'inverse, exposer les cellules bêta à l'interféron-alpha augmente fortement leur visibilité. La quantité de complexes pHLA présentée est plus importante que dans des cellules non stimulées, les peptides restreint aux allèles HLA-B sont enrichis et la diversité des peptides est accrue. Certains neo-épitopes ont été trouvés, résultant d'épissage alternative au niveau de l'ARN, au niveau peptidique ou de modifications post-traductionnelles. Le protéasome et l'immuno-protéasome ont été identifiés comme des acteurs clés dans la formation d'une grande partie de ces séquences.

Ces résultats dressent une cartographie exhaustive des cibles antigéniques exposées par les cellules bêta pour la reconnaissance de la part des lymphocytes T auto-immuns. Ces antigènes offrent la possibilité de suivre les réponses auto-immunes et anti-CVB au cours de l'histoire naturel du DT1.

Mots-clés : Diabète de type 1, complexe majeur d'histocompatibilité, immunopeptidomique, auto-immunité, Coxsackievirus B, interféron alpha.

Résumé substantiel

La visibilité antigénique de la cellule β contribue-t-elle au diabète de type 1 ? Analyse par peptidomique HLA

Le diabète de type 1 (DT1) est une maladie auto-immune, qui résulte de la destruction des cellules β, productrices d'insuline au sein des îlots de Langerhans situés dans le pancréas. Le DT1 apparait généralement chez les enfants avec un pic d'incidence entre 10 et 14 ans, mais il peut survenir à tous les âges. L'incidence du DT1 varie selon les régions géographiques mais augmente de façon constante depuis plusieurs années. La déclaration clinique du DT1 survient souvent après une longue phase asymptomatique caractérisée par l'infiltration du pancréas par des cellules immunitaires (insulite). Cet infiltrat mène à la destruction progressive des cellules β, ce qui se traduit par une augmentation de la glycémie. Une fois cliniquement déclaré, le principal traitement du DT1 consiste en l'injection d'insuline exogène plusieurs fois par jour, à vie. Ce traitement, essentiellement palliatif, est contraignant et ne permet pas d'empêcher certaines complications dégénératives telles que les rétinopathies, néphropathies et neuropathies.

Si l'étiologie du DT1 n'est pas encore clairement déterminée, plusieurs facteurs de risque du DT1 ont été identifiés. Premièrement, plusieurs gènes sont associés avec le développement de la maladie, le principal étant le gène du complexe majeur d'histocompatibilité (CMH, aussi appelé HLA) de classe II (HLA-II). Certains haplotypes sont en effet considérés comme protecteurs alors que d'autres indiquent une susceptibilité génétique. Dans une moindre mesure, les gènes du HLA de classe I sont également associés avec le DT1. Deuxièmement, de nombreux facteurs d'origine environnementale ont été discutés et le seul qui soit accepté largement est un Entérovirus, le Coxsackievirus B (CVB). Bien qu'une relation causale manque entre le DT1 et l'infection par CVB, qui est un virus très répanduchez les jeunes enfants, responsable de syndromes pseudo-grippaux, de multiples études ont mis en évidence un lien de corrélation. En effet, une association temporelle a pu être documentée entre la détection d'ARN viral dans les selles ou la présence d'anticorps neutralisant dirigés contre CVB et la progression vers le DT1. De plus, les protéines virales ont été retrouvées de façon plus fréquente sur des coupes histologiques de pancréas de donneurs DT1 en comparaison à celles de donneurs non diabétiques, ce qui suggère aussi une corrélation spatiale.

La diversité et la nature des cellules immunitaires qui infiltrent le pancréas jouent un rôle clé dans le développement du DT1, puisque les cytokines qu'elles produisent vont façonner la réponse immunitaire. Chronologiquement, la première cytokine observée avant même le déclenchement de la maladie, dans le sang d'enfants à risque, est l'interféron (IFN)-α. Cette cytokine est aussi impliquée dans des mécanismes immunitaires antiviraux et permet de limiter la réplication et la propagation virale.

Les cellules effectrices reconnues comme destructrices des cellules β sont les lymphocytes T (LT) $CD8^+$ cytotoxiques, qui reconnaissent les complexes peptides-HLA de classe I (pHLA-I) exprimés à la surface des cellules β . De façon surprenante, notre laboratoire a montré que leur fréquence est similaire dans le sang de donneurs diabétiques et sains, et leur phénotype est majoritairement naïf dans les deux groupes. Ces résultats suggèrent qu'un état universel d'auto-immunité « bénigne » existe, et interroge sur les mécanismes qui gouvernent le déclenchement du DT1. Ce changement d'état pourrait dépendre des cellules β elles-mêmes, qui deviendraient davantage visibles et vulnérables face au système immunitaire quand elles sont exposées à des conditions de stress. De plus, plusieurs autres protéines contenues dans les granules d'insuline sont sécrétées avec l'insuline elle-même, sous forme soluble ou dans des

vésicules extracellulaires, exosomes, granules entiers. Ces composants peuvent être phagocytés, apprêtés et présentés par des cellules dendritiques dans les ganglions pancréatiques, et ainsi mener à l'activation des LT auto-réactifs.

Mes travaux de thèse ont donc examiné la visibilité de la cellule β soumise à différents stress externes, en caractérisant les peptides présentés à sa surface au sein des complexes pHLA-I (immunopeptidome) et en évaluant la contribution de certains acteurs moléculaires à leur formation. Plus précisément, en utilisant une lignée de cellules β humaines et des îlots pancréatiques humains isolés, je me suis intéressée à l'infection par CVB, ainsi qu'à l'effet de l'IFN-α.

1. La visibilité de la cellule β face aux mécanismes d'échappement du virus

Dans un premier temps, l'infection d'une lignée cellulaire β avec différents sérotypes de CVB et l'utilisation de techniques de peptidomique-HLA, qui consistent à immunoprécipiter les complexes pHLA exposés à la surface des cellules pour ensuite séquencer les peptides par spectrométrie de masse, a révélé une présentation réduite de peptides viraux. De façon cohérente, ce phénomène s'accompagne d'une réduction de l'expression de molécules HLA-I membranaires. Ces observations résultent très certainement des différents mécanismes viraux mis en place lors de l'infection pour éviter de déclencher une réponse immunitaire antivirale. Cette approche a également été complémentée par des prédictions *in silico*, qui ont permis l'identification de candidats supplémentaires.

Afin de s'assurer de la pertinence immunologique des séquences trouvées, celles prédites comme se liant aux allèles HLA-A*02:01 (A2) et A*03:01 (A3), qui sont exprimés par le modèle cellulaire choisi et qui sont les allèles les plus fréquents dans la population caucasienne, ont été étudiées pour leur reconnaissance par des LT $CD8^+$. Ces LT sont repérés grâce à des multimères HLA-I chargés avec les peptides d'intérêt. L'analyse de LT issus du sang d'adultes et d'enfants sains et diabétiques de type 1 a montré que seule une partie des peptides viraux sont reconnus et que la fréquence de ces LT anti-CVB est basse, et similaires entre individus sains ou diabétiques. De plus, le phénotype de ces LT est souvent majoritairement naïf, ce qui suggère une faible réponse LT $CD8^+$. L'étude de splénocytes ainsi que de LT des ganglions lymphatiques pancréatiques a néanmoins démontré que les cellules issues de cet organe arborent un phénotype effecteur/mémoire. Ces observations indiquent que les cellules circulantes ne sont pas toujours un très bon miroir de la réponse antivirale qui a lieu dans le tissu et appellent des études supplémentaires dans le pancréas et les ganglions drainants.

Même si peu de peptides viraux sont présentés à la surface des cellules β infectées, nous avons pu montrer que d'un point de vue quantitatif, cette présentation est suffisante pour permettre une lyse par les LT $CD8^+$. Cependant, nos travaux ont aussi souligné que les LT $CD8^+$ ne sont pas les principaux instigateurs de la destruction des cellules β, qui est au contraire majoritairement due au cycle lytique du CVB. Ces résultats soutiennent l'idée qu'un vaccin anti-CVB pourrait aider à augmenter la réponse immunitaire contre ce virus et ainsi à prévenir le développement du DT1. Des études cliniques sont en cours pour évaluer la sécurité et les bénéfices d'un vaccin pentavalent. Les épitopes que nous avons répertoriés associés à des multimères HLA peuvent permettre de suivre la réponse immunitaire liée au vaccin tout en distinguant celle issue d'une infection naturelle. En effet, le vaccin ne contient que des protéines structurales du virus alors que la majorité des peptides retrouvés après infection dérivent de protéines non-structurales.

Enfin, nous avons pu mettre en évidence une reconnaissance croisée entre un épitope viral issu d'une protéine non-structurale et un antigène endogène issu de GAD2, qui n'est pas présenté par les cellules β infectées mais l'est à un état basal. Cette observation peut être

d'importance pour mieux comprendre les mécanismes du développement du DT1 après une infection par CVB et nécessite des études approfondies. Elle ne devrait pas déclencher de réponse auto-immune post-vaccination puisque le vaccin ne contient que des protéines structurales.

2. Effets de l'IFN-α

La deuxième partie de mon travail de thèse a porté sur l'étude de l'effet de l'IFN-α sur la visibilité antigénique des cellules β. En utilisant le même modèle β cellulaire et une approche de peptidomique-HLA, nous avons montré que l'IFN-α modifie l'immunopeptidome à plusieurs niveaux.

Ainsi, l'IFN-α augmente l'expression de molécules HLA-I à la surface des cellules, ce qui corrèle avec un nombre plus élevé de peptides. Ces peptides sont en moyenne plus abondants que ceux retrouvés sur des cellules non traitées. Le développement d'une analyse bio-informatique spécifique a permis d'identifier des peptides conventionnels, des peptides portant des modifications post-traductionnelles (PTMs), des peptides dérivés de l'épissage alternatif au niveau de l'ARNm ou d'épissage peptidique. Des stratégies de correspondance de spectres ont mené à la validation de 4 types de PTMs retrouvées dans 12 séquences différentes. De la même façon, 1 peptide issu de l'épissage alternatif d'ARNm d'insuline a pu être validé. Bien qu'elles nécessitent une étude mécanistique afin de corroborer leur origine protéasomale, 11 séquences résultant d'épissage peptidique ont été confirmées.

Exposer les cellules β à l'IFN-α accroit le nombre de protéines sources des peptides présentés par des complexes pHLA, qui sont autrement absentes de la surface cellulaire, reflétant le stress accru subit par le réticulum endoplasmique. En parallèle, le nombre de peptides dérivant de protéines des granules sécrétoires diminue. Malgré cela, de nouveaux antigènes issus des granules d'insuline ont été identifiés. L'IFN-α augmente donc la diversité et la quantité de protéines représentées à la surface cellulaire. Il accroit aussi singulièrement le nombre de peptides prédits comme étant des ligands des allèles HLA-B, de sorte que ce mécanisme soit le principal contributeur à l'augmentation générale en molécules HLA-I. Nous avons pu montrer que ce phénomène est dû à l'expression des gènes HLA-B qui croît plus fortement que celle des allèles HLA-A ou -C. La réponse des LT à ces séquences doit encore être caractérisée, mais cette découverte remet en question l'emphase actuelle qui est faite sur les ligands HLA-A alors que le facteur de risque génétique ainsi que les altérations principales de l'immunopeptidome de la cellule β exposée à des stress inflammatoires semblent davantage axés vers les allèles HLA-B. Nos résultats exhortent à mettre les ligands HLA-B au centre de l'attention dans le DT1.

Curieusement, nous avons également pu montrer que l'exposition à cette cytokine induit la présentation de peptides restreint à l'HLA-E, qui est un allèle impliqué dans la tolérance par ses interactions avec les cellules NK et LT $CD8^+$. HLA-E présente un intérêt grandissant dans le contexte du DT1, puisque son expression accrue pourrait représenter un mécanisme de défense pour limiter la vulnérabilité de la cellule β.

Par ailleurs, en employant des inhibiteurs du protéasome et de l'immuno-protéasome, qui sont des composant centraux de la voie d'apprêtement et de présentation des antigènes, nous avons pu mettre en évidence des groupements de peptides qui dépendent de l'expression de l'un de ces complexes enzymatiques, ou des deux. Ces résultats seront à mettre en lien avec la réponse des LT à ces séquences, mais ont un potentiel thérapeutique étant donné que plusieurs inhibiteurs du protéasome et de l'immuno-protéasome sont en développement et en utilisation clinique.

Enfin, en nous concentrant sur un sous-ensemble des ligands HLA-A2, nous avons constaté que les réponses des LT $CD8^+$ de donneurs sains sont peu fréquentes et essentiellement naïves. Bien qu'attendues, ces observations nous ont permis de sélectionner les peptides qui seront étudiés plus en profondeur dans des expériences de confirmation, qui compareront des donneurs sains et diabétiques de type 1, adultes et enfants. La même approche sera utilisée pour caractériser la réponse aux peptides restreint à HLA-A3 et aux allèles HLA-B.

En résumé, les données accumulées confirment que CVB détourne la machinerie cellulaire de son hôte afin de prévenir sa détection par le système immunitaire, ce qui a été appuyé par la faible réponse LT $CD8^+$ que nous avons observée. Par conséquent, l'immunopeptidome endogène de la cellule β est moins exposé. A l'inverse, l'IFN-α enrichie l'immunopeptidome en favorisant la diversité des peptides, en augmentant leur quantité, et en facilitant la formation de néo-épitopes. De plus, cette cytokine accroit la quantité de molécules HLA-I à la surface des cellules. Puisque la production d'IFN-α peut découler de l'infection virale, ces travaux indiquent collectivement que la grande variation dans l'immunpeptidome de la cellule β est un élément clé qui peut expliquer le changement d'un état d'auto-immunité bénigne vers le développement du DT1.

En conclusion, ce travail dresse une cartographie exhaustive des cibles antigéniques exposées par les cellules β pour la reconnaissance par des lymphocytes T auto-immuns. Ces antigènes offrent la possibilité de suivre les réponses auto-immunes et anti-CVB au cours de l'histoire naturelle du DT1 et fournit des pistes à explorer pour comprendre le lien entre l'infection par CVB, l'inflammation qui en résulte et le DT1.

Acknowledgments

« We pass through the present with our eyes blindfolded. We are permitted merely to sense and guess at what we are actually experiencing. Only later when the cloth is untied can we glance at the past and find out what we have experienced and what meaning it has. »

Milan Kundera

As far as I can remember, I have always been willing to do a PhD. But little did I know how unusual and intense this experience would be. I believe that unless you do one yourself, it's difficult to appreciate what a challenge it is. But you do not need to understand one's journey to support them and I feel extremely lucky because my path has crossed that of formidable people through this otherwise lonely quest. I hope that the few upcoming words will translate properly the limitless gratitude I feel toward all of you!

First, I would like to thank the members of my thesis committee, for dedicating a few hours of their certainly busy agendas to examine my work, allowing me to improve it, and fueling a discussion that I don't doubt will be very enriching. Thank you Prof. Miriam Cnop and Dr. Wei Wu for reviewing my manuscript and Dr. Loredana Saveanu, Prof. David Price and Prof. Didier Hober for attending my defense.

Roberto, we have had a few open conversations over the years and I am sure you already know how immensely grateful I am. You have always been supportive of me, even when that meant going out of your comfort zone or allowing me to check whether the grass was greener elsewhere. You have encouraged me to follow my ambitions and to take my own path while guiding me through. As busy as you could be, you always had time to listen and discuss. You never made me feel inferior, despite the fact that you are obviously more knowledgeable than I am, and instead simply questioned my ideas so that I would gain confidence and perfect my thinking process. Humanely and scientifically, you've been an ideal mentor. Thank you for all the support you provided, the trust you gave, the involvement you showed.

Sylvaine, even if we have not interacted as much, you've been very supportive and have helped me grow as a scientist. It was a true pleasure to be part of your team and I knew I could rely on you if needed, thank you. To be honest, one of the biggest strengths that you both have is your awareness that life happens, and sometimes science has to come second, not first. This

is exceptionally precious and I am grateful that I had the chance to perform my PhD in such an environment.

I would also like to express my deepest appreciation to Nicola, you have shared so much of your expertise and I have learned a lot in the last few years. Your approach to science has greatly balanced the one that Roberto has and I have found my own somewhere in the middle. Thank you for all of that, and thank you for your joy and excitement about the slightest result. It was truly refreshing!

To all my fellow lab mates: a big thank you! Thank you for your patience in listening to my quite foreign projects. I know all the great scientific discussions we had required a specific effort from you – sorry not sorry, I don't work on T cells! Thank you also for the emotional and psychological comfort, for the chit-chatting, and for all the moments (and all the food and drinks) that we have shared over the years.

Javi, you arrived in the lab not so long after me, and gosh you've been annoying since then. Your humor is often misunderstood, and you are even more rigid than I am. So, thanks for allowing me to take some perspective on my own inflexibility – relax, I'm joking! A bit more seriously, thank you for all our conversations, whether they were about science, politics, movies/TV shows, or our lives, they each taught me something and you were often providing good advice.

P.S: you should be thanking me for tolerating you in MY office for so long. I am now handing it to you, so please take care of it.

Barbara, you are a rare kind of lab mate, and… it's maybe better that way. You're a digger of the weirdest songs, and you know by heart lines of so many TV shows and movies. I am glad that there is only one of you on our team. I am not sure my (lack of) pop culture could handle more than one. I'm joking but truly, we need people like you. You dedicate so much of your time to others and you have helped me on so many levels. I have lost count of how many times you've listened to me complaining about one thing or another (let's admit it, often about Javi), or how often I come to you to find stuff in the lab, even though I'm here for longer than you, or how we would sometimes leave more than one hour after finishing our day because we would just talk forever – we should have definitely gone to a bar in these instances. Oh! I almost forgot but all my friends (yes, Rémy, I'm speaking in your name specifically) thank you for sharing your cookie recipe, it's been put to good use!

Clémentine, you're one of the greatest people I've met. You are incredibly knowledgeable yet very humble. I genuinely enjoyed working with you and I wish we had more opportunities to do so. Thank you also for our deep philosophical discussions, even if I'm not sure the beta cells appreciated them as much as I did, it helped me navigate life. And thank you for letting me inflict on you that you would be my co-*correspondant informatique* for a few months, before taking it over on your own. Now I can only wish you a PhD experience as happy as mine was!

Fatou, thank you for making me realize how funny I am. Surprisingly you're the only one aware of it… (and it's definitely not related to how easy it is to make you laugh). Your sympathetic ear and your constant good mood were always joyful breaks that my heart needed. I will miss hearing you laughing from two floors below.

Fede, my co-PhD, my Chouffe companion but mostly, the one that we never know if or when she comes in. You've been so easy to talk to and you're also a lot of fun and I'm glad that I got the chance to enjoy this side of you! And Orlando, one of the kindest souls ever. You're the two people in this lab with whom I've truly shared a project and I know for sure that those projects are in good hands. Keep up to good work, and good luck to both of you!

Anna, I'm so grateful that I had the chance to visit you in Italy. Your hospitality and generosity were beyond all my expectations and I truly relished our late-night conversations. Your home is a haven of peace and I hope that I'll be able to come again soon.

Widad and Yohan, we haven't really worked together, but you're both great people… except for the fact that I feel old near you. Yohan thank you for your jokes, I might have found someone to my level (I'm not sure it's a compliment, maybe ask Fatou)! Widad thank you for your kindness and friendliness, I hope your ambitions will bring you where you want.

The day I'm writing this, Gigi, you're still officially part of our lab. I will not put you in the section of the members that have left (I'm not an expert, but I think it's called denial), you have spent so much time here, and given so much, that you will always be part of the team. When I first started in the lab, I ended up sitting next to you and you had that calming aura around you. I felt welcomed right away and soon enough and surely like everybody else, I considered you as our lab mom. There is truly no good depiction of how kindhearted you are. You're a gem.

I would also like to thank all the other members or ex-members of our lab: Abdou, Zhicheng, Asia, Vanesa, Virginie, Sarah, Aurore, Julien, Noémie, Camille, Stefano, Priscilla. I have also had the pleasure to work with all the great people of our beloved Cassini building. A special thank you to Latif, for his time and constructive discussions, always questioning my procedures. You've cost me a lot in cakes but I'll admit that it was worth it, your opinion was helpful. Maybe I should have asked for the same payment type for the many questions you had about mass spec…

And to all the ones that have already left, Claire D., some would say that you are too rigid, too much about the rules, and overall a bit annoying. But that's exactly why I'm grateful that you ended up in our lab. First, we needed that, and then it made me feel less alone trying to get the rules respected. You have helped me quite a few times, giving your very insightful suggestions and bits of advice. You're someone I look up to and I hope that one day I'll be able to guide inexperienced students as you've guided me.

Lucas and Sherlina, thank you for your technical help and for alleviating parts of my burden when I was the most desperate!

Claire M. thank you for training me when I arrived in the lab and for your daily help and our conversations. You were so hard-working, devoted, and willing to help even if that meant having looooong days. Believe me, from the day you left the lab to the one I will be leaving, I've missed you. You're just great, and I hope you'll realize to what extent one day.

Boban, you are such a character, there is no other way to put it. The day I arrived you made me uncomfortable (but in a weirdly good way) and you made me laugh quite a lot. I also admire your open-mindedness and I would like to thank you for making me feel welcome!

Mahmoud thank you for being you. You're in the same category as Boban. You too are a character and you've made me laugh a few times, but also reflect on life and on what's important. MEA thank you for your refreshing candor and kindness.

To the ones in Oxford,

I could not have done all this work without the help of Nicola's group members. Rob, I doubt you remember that, but it's one of our conversations at the first HUPO-HIPP that convinced me to chat with Nicola and see whether there was room to set up a collaboration. Thank you for that, for running most if not all of my samples, and for teaching me some tricky data analyses, like DIA (or should I say DIE?). Your endless jokes, cheerfulness, and bits of

drama have made my visits incredibly pleasant. You're a beautiful soul and I'm grateful for all the moments we shared. Annalisa, it was so easy to learn with you! Thank you for being so sweet, helpful, and understanding. Hanqing, thank you for encouraging me to go back into Python, and of course for all your bioinformatics help all along. Rebecca and Wayne, thank you for sharing your knowledge with me and for our very interesting conversations. Arun, the social one, thanks for bringing everybody together and always pushing for outings. To you and all the others, Thierry, Isaac, Nikki, Xu, thank you for welcoming me on each of my visits, it was a delight!

To the medical staff and the donors,

I have had a hard time recruiting blood donors. I would like to acknowledge them and the nurses of the Copernic building for drawing them and for going out of their comfort zone.

To my loved ones,

Mom, Dad, you've both always supported me in your own way. Mom, the values you've instilled in me have driven me where I am: with hard work and persistence nothing's impossible. Dad, your constant worry that I work too much actually forced me to take time off. "Never get so busy making a living that you forget to make a life". Thank you both for believing in me.

Parrain, Isabelle, Nathalie, Caroline, and Jacques, thank you for our many shared moments (let's be honest, often around good food and wine). You have invariably shown your interest and it warms my heart to feel so much cared for. I appreciate so profoundly all your little gestures and consideration.

I am not one of many friends, but I believe that I am one of genuine friendships. Thank you, Julie, Marie, Gillou, Yann, Mélo, the JAS, Anto, Alex, Guigui, Adrien, Edwige for keeping me afloat, for reminding me that there is more to life than just work, for all the drinks, the parties, the weekends and so much more. I would have drowned without you all!

Gaby. There are no words to describe who you are to me and only a few people could possibly understand it. You have passed much of your life wisdom on me. I'm eternally grateful to you because if I was so well equipped to do a PhD, it's largely thanks to you.

Boulie, you're the sister that I have never had and so much more than that. You've made me a godmother to the sweetest little girl possible – her candor and seeing life through her lens made me realize that my worries were not that important after all. You've been the one constant

in my life for nearly two decades now. You have listened to me passionately trying to explain what I do for a living for hours at a time, without complaining. You have continually supported me, comforted me, and always made me confident that I could do it when I doubted myself. I could write dozens of pages, yet words could not express my gratitude to you.

Lastly, Ludo. You're probably the one that would deserve an acknowledgment book or a prize of some sort. You put up with me every single day, from the beginning to the end of my thesis. You endured all the emotions I was going through, worse, you shared them - whether it was the disappointment of an experiment not working, the frustration of being stuck with things out of my reach, or the overwhelming enthusiasm when the results were excellent. You patiently listened with interest to the details of my project and as if that wasn't enough, you also allowed me to take it to the next level. You brainstormed with me and spent entire evenings helping me debug my scripts after you had finished your own day's work. You taught me what you knew, even if I have to admit that it took me a while to do these unit tests, if they are there it's because of you. You were one of the only ones to question what I was doing, with your naive but often relevant and constructive eye, and with a benevolence that is beyond understanding and that represents so much of who you are. In the end, this thesis, is also a bit yours. In the most intense periods for me, you ensured my survival, you took it upon yourself without complaining to manage our daily life, and not even for a second you made me feel guilty about it. You also pulled me out of work when you felt I was reaching my limits. You pushed me to go out, to see something else, to go away for the weekend. You were on all fronts, and never, ever, did I expect so much from a partner. I am eternally grateful for all you have done, and I hope that when you will need my full support, I can live up to what you have done for me. We've survived a thesis, a pandemic, a quarantine. Now I just can't wait to see where life will take us and what our future will be like. I love you.

And thank **you**, reader, for making it so far in the many pages of gratitude that I wrote and of which you are now a part.

As this is probably the only part of my manuscript that most of my relatives will be interested in and willing to read, I hope you will understand that I needed to translate it into French for some of them. Thank you DeepL for making this step faster and easier.

Version française

« Nous traversons le présent les yeux bandés. Tout au plus pouvons-nous pressentir et deviner ce que nous sommes en train de vivre. Plus tard seulement, quand est dénoué le bandeau et que nous examinons le passé, nous nous rendons compte de ce que nous avons vécu et nous en comprenons le sens. »

Milan Kundera

D'aussi loin que je me souvienne, j'ai toujours eu envie de faire un doctorat. Mais j'étais loin de me douter à quel point cette expérience serait inhabituelle et intense. Je pense qu'à moins d'en faire un soi-même, il est difficile d'apprécier le défi que cela représente. Mais il n'est pas nécessaire de comprendre le parcours d'une personne pour la soutenir et je me sens extrêmement chanceuse car mon chemin a croisé celui de personnes formidables à travers cette quête autrement solitaire. J'espère que les quelques mots à venir traduiront correctement la gratitude sans limite que je ressens envers vous tous !

Tout d'abord, je voudrais remercier les membres de mon comité de thèse, pour avoir consacré quelques heures de leurs agendas certainement très chargés pour examiner mon travail, me permettre de l'améliorer, et alimenter une discussion qui, je n'en doute pas, sera très enrichissante. Merci à Prof. Miriam Cnop et Dr. Wei Wu pour avoir revu mon manuscrit et à Dr. Loredana Saveanu, Prof. David Price et Prof. Didier Hober pour avoir assisté à ma soutenance.

Roberto, nous avons eu quelques conversations ouvertes au fil des ans et je suis sûre que tu sais déjà à quel point je te suis immensément reconnaissante. Tu m'as toujours soutenu, même lorsque cela signifiait sortir de ta zone de confort ou me permettre de vérifier si l'herbe était plus verte ailleurs. Tu m’as encouragé à suivre mes ambitions et ma propre voie tout en me guidant sur le chemin. Aussi occupé que tu aies pu être, tu avais toujours le temps de m’écouter. Tu ne m’as jamais fait sentir que j'étais inférieure, bien que tu en saches manifestement plus que moi, et au contraire, tu as simplement remis en question mes idées afin que je prenne confiance et que je perfectionne mon processus de réflexion. Sur le plan humain et scientifique, tu as été un mentor idéal. Merci pour tout le soutien que tu m’as apporté, la confiance que tu m’as donnée et pour l'implication dont tu as fait preuve.

Sylvaine, même si nous avons moins échangé, tu m'as beaucoup soutenu et tu m'as aidé à grandir en tant que scientifique. C'était un vrai plaisir de faire partie de ton équipe et j'ai toujours su que je pouvais compter sur toi si besoin, merci. Pour être honnête, l'une des plus grandes forces que vous possédez tous les deux, est que vous soyez conscient que la vie ne se résume pas qu'à la science et que parfois la science doit passer en second, et non en premier. C'est exceptionnellement précieux et je suis reconnaissante d'avoir eu la chance de réaliser mon doctorat dans un tel environnement.

J'aimerais également exprimer ma profonde gratitude à Nicola. Tu as partagé une grande partie de ton expertise et j'ai beaucoup appris au cours de ces dernières années. Ton approche de la science a grandement équilibré celle de Roberto et j'ai trouvé la mienne quelque part au milieu. Merci pour tout cela, et merci pour ta joie et ton excitation au moindre résultat. C'était vraiment rafraîchissant !

À tous mes collègues de laboratoire : un grand merci ! Merci d'avoir eu la patience d'écouter mes projets « différents » pour vous. Je sais que toutes les grandes discussions scientifiques que nous avons eues vous ont demandé un effort particulier – déso, pas déso, je ne travaille pas sur les lymphocytes T ! Merci aussi pour le réconfort émotionnel et psychologique, pour les bavardages et pour tous les moments (et toute la nourriture et tous les verres) que nous avons partagés au fil des années.

Javi, tu es arrivé au labo peu de temps après moi, et depuis, tu n'as cessé d'être agaçant. Ton humour est souvent mal compris, et tu es encore plus rigide que moi. Alors, merci de me permettre de prendre du recul sur ma propre rigidité – détends toi, je plaisante ! Un peu plus sérieusement, merci pour toutes nos conversations, qu'elles aient porté sur la science, la politique, les films/émissions de télévision ou nos vies, elles m'ont toutes appris quelque chose et tu as souvent été de bon conseil.

P.S : tu devrais me remercier de t'avoir toléré dans MON bureau pendant si longtemps. Je te le remets maintenant, alors prends-en soin.

Barbara, tu es d'un genre rare, et... c'est peut-être mieux ainsi. Tu es une dénicheuse des chansons les plus étranges, et tu connais par cœur les répliques de tellement de séries et de films. Je suis contente qu'il n'y en ait qu'une comme toi dans l'équipe. Je ne suis pas sûre que ma culture pop (ou plutôt absence de) aurait pu endurer plus d'une Barbara. Je rigole, mais sincèrement, on a besoin de gens comme toi. Tu consacres tellement de temps aux autres et tu m'as aidé à bien des égards. Je ne compte plus le nombre de fois où tu m'as écouté me plaindre

d'une chose ou d'une autre (admettons-le, souvent à propos de Javi), ou le nombre de fois où je viens te voir pour trouver des trucs dans le labo, même si je suis ici depuis plus longtemps que toi, ou comment on partait parfois plus d'une heure après avoir fini notre journée parce qu'on papotait - on aurait vraiment dû aller dans un bar dans ces cas-là. Oh ! j'allais oublier mais tous mes amis (oui, Rémy, je parle en ton nom précisément) te remercient d'avoir partagé ta recette de cookies, elle a été utilisée à bon escient !

Clémentine, tu es l'une des personnes les plus formidables que j'ai rencontrées. Tu es incroyablement compétente et pourtant très humble. J'ai vraiment apprécié de travailler avec toi et j'aurais adoré avoir davantage d'occasions. Merci aussi pour nos profondes discussions philosophiques, même si je ne suis pas sûre que les cellules bêta les aient appréciées autant que moi, cela m'a aidé à avancer. Et merci de m'avoir laissé t'infliger le fait que tu serais ma co-correspondante informatique pendant quelques mois, avant de l'être toute seule. Je te souhaite une thèse aussi heureuse que la mienne a pu l'être !

Fatou, merci de m'avoir fait réaliser à quel point je suis drôle. Étonnamment, tu es la seule à en être consciente... (et ça n'a sûrement rien à voir avec le fait que tu sois bon public). Ton oreille compatissante et ta bonne humeur constante ont toujours été des pauses joyeuses dont j'avais besoin. Cela va me manquer de ne plus t'entendre rire depuis deux étages plus bas.

Fede, ma co-doctorante, ma compagne de Chouffe mais surtout, celle dont on ne sait jamais si ou quand elle vient au labo. Tu es quelqu'un à qui il est facile de parler et tu es aussi très drôle. Je suis heureuse d'avoir eu la chance de profiter de ce côté-là de toi ! Et Orlando, l'une des personnes les plus gentilles qui soient. Tous les deux, vous êtes les deux personnes de ce laboratoire avec qui j'ai vraiment partagé un projet et je sais avec certitude que ces projets sont entre de bonnes mains. Bonne chance à tous les deux !

Anna, je suis tellement reconnaissante d'avoir eu la chance de te rendre visite en Italie. Ton hospitalité et ta générosité ont dépassé toutes mes attentes et j'ai vraiment apprécié nos conversations nocturnes. Ta maison est un havre de paix et j'espère que je pourrai revenir prochainement.

Widad et Yohan, nous n'avons pas vraiment travaillé ensemble, mais vous êtes tous les deux des gens géniaux... à part le fait que je me sens vieux près de vous. Yohan merci pour tes blagues, j'ai peut-être trouvé quelqu'un à mon niveau (je ne suis pas sûre que ça soit un compliment, demande peut-être à Fatou) ! Widad merci pour ta gentillesse, j'espère que tes ambitions t'emmèneront là où tu veux.

Le jour où j'écris ces lignes, Gigi, tu fais toujours officiellement partie du labo. Je ne te mettrai pas dans la section des membres qui sont partis (je ne suis pas experte, mais je crois que ça s'appelle le déni), tu as passé tellement de temps ici, et donné tellement, que tu feras toujours partie de l'équipe. Lorsque j'ai commencé dans le laboratoire, je me suis retrouvée assise à côté de toi et tu avais cette aura apaisante autour de toi. Je me suis sentie tout de suite la bienvenue et très vite, comme tout le monde, je t'ai considérée comme notre maman de labo. Il n'y a vraiment pas de bonne description de la gentillesse dont tu fais preuve. Tu es une perle.

Je voudrais également remercier tous les autres membres ou ex-membres de notre laboratoire : Abdou, Zhicheng, Asia, Vanesa, Virginie, Sarah, Aurore, Julien, Noémie, Camille, Stefano, Priscilla. J'ai également eu le plaisir de travailler avec toutes les personnes formidables de notre cher bâtiment Cassini. Un merci particulier à Latif, pour son temps et ses discussions constructives, remettant toujours en question mes procédures. Tu m'as coûté beaucoup en gâteaux mais j'avoue que cela en valait la peine, ton avis m'a été utile. J'aurais peut-être dû demander le même type de paiement pour les nombreuses questions que tu m'as posé sur la spectro de masse...

Et à tous ceux qui sont déjà partis, Claire D., certains diront que tu es trop rigide, trop à cheval sur les règles, et globalement un peu chiante. Mais c'est exactement pour ça que je suis reconnaissante que tu aies atterri dans notre laboratoire. D'abord, nous en avions besoin, et ensuite cela m'a permis de me sentir moins seule à essayer de faire respecter les règles. Tu m'as aidé à plusieurs reprises, en m'offrant des suggestions et des conseils très perspicaces. Tu es quelqu'un que j'admire et j'espère qu'un jour je pourrai accompagner des étudiants inexpérimentés comme tu l'as fait pour moi.

Lucas et Sherlina, merci pour votre aide technique et pour avoir allégé une partie de mon fardeau lorsque j'étais le plus désespérée !

Claire M., merci de m'avoir formée à mon arrivée au laboratoire et pour ton aide quotidienne et nos conversations. Tu étais volontaire, dévouée et prête à aider même si cela impliquait d'avoir des journées très longues. Crois-moi, du jour où tu as quitté le laboratoire à celui où je vais le quitter, tu m'as manqué. Tu es tout simplement géniale, et j'espère que tu réaliseras un jour à quel point.

Boban, tu es un tel personnage, il n'y a pas d'autre façon de le dire. Le jour de mon arrivée, tu m'as mise mal à l'aise (mais d'une manière étrangement sympathique) et tu m'as fait

beaucoup rire. J'admire également ton ouverture d'esprit et je tiens à te remercier de m'avoir fait sentir la bienvenue !

Mahmoud, merci d'être toi. Tu es dans la même catégorie que Boban. Toi aussi tu es un personnage et tu m'as fait rire plusieurs fois, mais aussi réfléchir sur la vie et ce qui est important. MEA, merci pour votre ta candeur rafraichissante et ta gentillesse.

À ceux d'Oxford,

Je n'aurais pas pu faire tout ce travail sans l'aide des membres du groupe de Nicola. Rob, je doute que tu t'en souviennes, mais c'est l'une de nos conversations lors du premier HUPO-HIPP qui m'a convaincue de discuter avec Nicola et de voir s'il était possible de mettre en place une collaboration. Merci pour cela, pour avoir lancé les acquisitions de la plupart, sinon la totalité, de mes échantillons, et pour m'avoir appris certaines analyses de données délicates, comme le DIA (ou devrais-je dire DIE ?). Tes plaisanteries sans fin, ta bonne humeur et ton sens du théâtre ont rendu mes visites incroyablement agréables. Tu es une belle âme et je te suis reconnaissant pour tous les moments que nous avons partagés. Annalisa, c'était si facile d'apprendre avec toi ! Merci d'être si douce, si serviable et si compréhensive. Hanqing, merci de m'avoir encouragé à me remettre à Python, et bien sûr pour toute l'aide que tu m'as apportée en bioinformatique. Rebecca et Wayne, merci d'avoir partagé vos connaissances avec moi et pour nos conversations intéressantes. Arun, le sociable du labo, merci d'avoir continuellement cherché à rassembler tout le monde et d'avoir poussé à faire des sorties. A toi et à tous les autres, Thierry, Isaac, Nikki, Xu, merci de m'avoir accueilli à chacune de mes visites, c'était un plaisir !

Au personnel médical et aux donneurs,

J'ai eu beaucoup de mal à recruter des donneurs de sang. Je tiens à les remercier ainsi que les infirmières du bâtiment Copernic pour les avoir prélevés et être sorties de leur zone de confort.

À mes proches,

Maman, papa, vous m'avez toujours soutenu à votre propre manière. Maman, les valeurs que tu m'as inculquées m'ont conduite là où je suis : avec du travail et de la persévérance, rien n'est impossible. Papa, ton souci constant de me voir trop travailler m'a obligé à prendre des pauses. "Ne sois jamais si occupé à gagner ta vie que tu en oublies de la vivre". Merci à tous les deux d'avoir cru en moi.

Parrain, Isabelle, Nathalie, Caroline, et Jacques, merci pour nos nombreux moments partagés (soyons honnêtes, souvent autour de bons plats et de bons vins). Vous avez invariablement manifesté votre intérêt et cela me fait chaud au cœur de me sentir si bien entourée. J'apprécie si profondément tous vos petits gestes et votre considération.

Je ne suis pas de celles qui a beaucoup d'amis, mais qui ont des amitiés sincères. Merci, Julie, Marie, Gillou, Yann, Mélo, le JAS, Anto, Alex, Guigui, Adrien, Edwige pour m'avoir maintenue à flot, pour m'avoir rappelé qu'il n'y a pas que le travail dans la vie, pour tous les verres, les fêtes, les week-ends et tant d'autres choses. Sans vous tous, je me serais noyée !

Gaby. Il n'y a pas de mots pour décrire qui tu es pour moi et seules quelques personnes pourraient le comprendre. Tu m'as transmis une grande partie de ta sagesse de vie. Je t'en serai éternellement reconnaissante car si j'étais si bien préparée pour faire un doctorat, c'est en grande partie grâce à toi.

Boulie, tu es la sœur que je n'ai jamais eue et tellement plus que cela. Tu as fait de moi la marraine de la plus adorable petite fille qui soit - sa candeur et voir la vie à travers ses yeux m'ont fait réaliser que mes soucis n'étaient pas si importants après tout. Tu es la seule constante dans ma vie depuis près de deux décennies maintenant. Tu m'as écouté passionnément essayer d'expliquer ce que je fais dans la vie pendant des heures, sans te plaindre. Tu m’as toujours soutenu, réconforté et fait croire en moi lorsque je doutais. Je pourrais écrire des dizaines de pages, mais les mots ne sauraient exprimer la gratitude que j’ai envers toi.

Ludo, tu es sûrement celui qui mériterait un livre entier de remerciements, ou un prix de quelque sorte qu’il soit. Tu m’as supporté chaque jour, du début à la fin de ma thèse. Tu as enduré toutes les émotions que je vivais, pire encore, tu les as partagés – que ce soit la déception d’une expérience qui ne marche pas, la frustration d’être bloquée par des choses hors de ma portée ou l’enthousiasme démesuré quand les résultats sont excellents. Tu m’as patiemment écouté avec intérêt raconter les détails de mon projet et comme si ça ne suffisait pas, tu m’as aussi permis de le faire progresser. Tu as brainstormé avec moi, tu as passé des soirées entières à m’aider à débuguer mes scripts, après avoir fini ta propre journée de travail. Tu m’as enseignée ce que tu savais, même si je dois admettre que j’ai mis du temps à les faire ces tests unitaires, s’ils sont là c’est grâce à toi. Tu as été un des seuls à remettre en question ce que je faisais, avec ton regard certes naïf mais souvent pertinent et constructif, et avec une bienveillance qui dépasse l’entendement et qui te caractérise tant. Cette thèse, au final c’est aussi un peu la tienne. Dans les périodes les plus intenses pour moi, tu as assuré ma survie, tu

as pris sur toi sans t'en plaindre la gestion du quotidien et pas une seconde tu ne m'as fait culpabiliser. Tu m'as aussi extirpé du travail quand tu sentais que j'atteignais mes limites. Tu m'as poussé à sortir, à voir autre chose, à partir en weekend. Tu as été sur tous les fronts, et jamais, vraiment jamais, je n'en attendais autant d'un partenaire. Je te suis éternellement reconnaissante pour tout ce que tu as fait, et j'espère que le jour où tu auras besoin de mon soutien intégral, je saurais me montrer à la hauteur de ce que tu as fait pour moi. On a survécu à une thèse, une pandémie, une quarantaine. Maintenant j'ai juste hâte de voir où la suite va nous mener et de quoi notre avenir sera fait. Je t'aime.

Et merci **lecteur**, d'avoir lu jusqu'au bout les mille et une page de remerciements que j'ai écrit et dont tu fais maintenant partie.

To my parents,

To Flo,

To Ludo,

« If you're following your calling, the fatigue will be easier to bear, the disappointments will be fuel, the highs will be like nothing you've ever felt. »

« Lorsque l'on suit sa vocation, la fatigue devient plus facile à supporter, les déceptions sont un carburant, et les réussites ont une saveur exceptionnelle. »

Phil Knight

Table of content

List of figures and tables 4
Abbreviations 5
Introduction 7
1 Type 1 diabetes 9
1.1 Epidemiology 9
1.2 Natural history 10
1.3 Etiology 11
1.3.1 Genetical factors 13
1.3.1.1 HLA genes 13
1.3.1.2 Other genes 13
1.3.2 Environmental factors 14
1.3.2.1 Enteroviruses 15
1.3.2.2 Microbiota 16
1.3.2.3 Hygiene hypothesis 16
1.4 Immunology of T1D 16
1.4.1 Islets immune infiltration 16
1.4.2 Cytokines 19
1.4.3 Humoral immune response: autoantibodies 20
1.4.4 Overview of T1D immunopathogenesis 21
1.5 T1D endotypes 22
2 The HLA system 24
2.1 HLA class I and class II 24
2.1.1 Gene structure, expression, and function 24
2.1.2 Classical and non-classical HLA-I molecules 25
2.1.3 Classical and non-classical HLA-II molecules 26
2.2 Antigen processing and presentation pathways 27
2.2.1 HLA-I antigen processing and presentation 28
2.2.1.1 Classical pathway 28
2.2.1.2 Alternative antigen sources and presentation pathways 29
2.2.1.2.1 Role of the immunoproteasome 29
2.2.1.2.2 Cross-presentation 30
2.2.1.2.3 Vesicular pathways 31

2.2.1.2.4 In β cells 32
2.2.2 HLA-II antigen processing and presentation 32
2.2.2.1 Classical pathway 32
2.2.2.2 Alternative antigen sources and presentation pathways 34
2.3 The immunopeptidome 35
2.3.1 Definition and identification 35
2.3.1.1 Binding prediction algorithms 35
2.3.1.2 Immunopeptidomics 36
2.3.1.2.1 Workflow 36
2.3.1.2.2 Type of acquisitions 37
2.3.1.2.3 Data analysis 37
2.3.2 Neo-antigens 38
2.3.2.1 Genomic mutations 38
2.3.2.2 Post-translational modifications 38
2.3.2.3 Peptide splicing 39
2.3.2.4 Other origins 40
2.3.3 T1D epitopes 41
3 Benign autoimmunity and loss of β-cell ignorance 42
3.1 We are all autoimmune 42
3.2 Accomplice β cells 45
3.2.1 Metabolic β-cell stress 45
3.2.2 β-cell susceptibility to its environment 46
3.2.3 β-cell antigenic visibility 47
3.2.4 β-cell dialogue with the immune system 49
3.2.5 β-cell targeting therapies 50
Aims of the PhD project 52
Results 53
Article #1 55
Article #2 109
Discussion and perspectives 181
Immune evasion of CVB-infected β cells 183
IFN-α increases the antigenic visibility of β cells 187
Technical limitations and improvements 191

Annexes 195

Annex 1: Review about the diabetogenic mechanisms of CVB (invited article for *Endocrine Reviews*) 197

Annex 2: Review on the evidence for the role of T cells in T1D 221

Annex 3: Review on β-cell-specific antigen processing and presentation pathways and implications in T1D 232

Annex 4: Original article showing that citrullinated GRP78 peptides can be recognized by circulating $CD8^+$ T cells 240

Annex 5: Original article on the benchmarking of a novel MARS algorithm to identify *cis*-spliced peptides 272

Annex 6: main Python pipeline developed for HLA-peptidomics analysis 310

References 315

List of figures and tables

Figure 1. T1D epidemiology 10
Figure 2. Natural history of T1D 12
Figure 3. Genes associated with T1D susceptibility 14
Figure 4. Insulitis evolution in T1D 17
Figure 5. Autoantibodies in T1D 20
Figure 6. Schematic representation of T1D immunopathogenesis 22
Figure 7. Location and organization of the HLA genes 24
Figure 8. MHC-I structure 25
Figure 9. MHC-II structure 27
Figure 10. The peptide cycle for HLA-I 28
Figure 11. Conventional HLA-I antigen processing and presentation pathway 29
Figure 12. Formation of the immunoproteasome 30
Figure 13. Antigens cross-presentation by dendritic cells 31
Figure 14. Recycling pathways in non-professional APCs 32
Figure 15. Conventional HLA-II antigen processing and presentation pathway 33
Figure 16. Workflow for the purification and identification of HLA-bound peptides by LC-MS/MS 36
Figure 17. Proteasome-catalyzed peptide splicing (PCPS) 40
Figure 18. Levels of evidence for islet-derived T-cell epitopes 41
Figure 19. Frequency of islet-reactive $CD8^+$ T cells in T1D patients and healthy individuals 43
Figure 20. β-cell biology weaknesses promoting autoimmune vulnerability 45
Figure 21. β-cell role in their own demise 46
Figure 22. Revised T1D immunopathogenesis 50
Figure 23. Cell and peptides numbers across experiments. 184
Figure 24. Abundance of INS_{15-24} (ALWGPGPAAA) and $GAD_{257-265}$ (KMFPEVKEK) peptides 186
Figure 25. Expression of HLA genes upon cytokine exposure 189
Figure 26. CVB and IFN-α mechanisms participating in triggering T1D autoimmunity 192

Table 1. Characteristics of T1D endotypes 23

Abbreviations

aAb	autoantibody
APC	antigen presenting cell
APP	antigen processing and presentation
CVB	Coxsackievirus B
DC	dendritic cell
DIPP	T1D Prediction and Prevention study
DiViD	Diabetes Virus Detection study
ER	endoplasmic reticulum
ERAAP /ERAP1	endoplasmic reticulum aminopeptidase
GAD	glutamic acid decarboxylase
GLP-1	glucagon-like peptide 1
GRP78	binding immunoglobulin protein
HbA1c	glycated hemoglobin A1c
HLA-I	human leukocyte antigen class I
HLA-II	human leukocyte antigen class II
IA-2	insulinoma antigen 2
ICI	insulin-containing islet
IDI	insulin-deficient islet
INS	insulin gene
IFN	interferon
IGRP	glucose-6-phosphatase 2
IL	interleukin
MMTT	mixed meal tolerance test
moDCs	monocyte-derived dendritic cells
MS	mass spectrometry
mTEC	medullary thymic epithelial cells
NOD	non-obese diabetic
nPOD	network for pancreatic organ donors with diabetes
ORF	open-reading frame
PBMC	peripheral blood mononuclear cells
pDC	plasmacytoid dendritic cell
pLN	pancreatic lymph node
TAP	transporter associated with antigen presentation
TEDDY	The Environmental Determinants of Diabetes in the Young
TNF	tumor necrosis factor
UPR	unfolded protein response
ZnT8	Zinc transporter 8

Introduction

1 Type 1 diabetes

Type 1 diabetes (T1D) is an autoimmune disease that results from the destruction of insulin-producing β cells within the islets of Langerhans in the pancreas[1]. This loss is driven by an intricate crosstalk of innate and adaptive immune cells, dominated by autoreactive effector $CD8^+$ T cells[2]. Endogenous insulin secretion becomes scarce and hyperglycemia develops. Among the early symptoms of the disease, the most common ones are polydipsia associated to polyuria and weight loss. T1D is clinically diagnosed upon measurements of glycemia >11.1 mmol/L, fasting glycemia >7mmol/L, or glycated hemoglobin A1c (HbA1c) >6.5%[3,4]. Currently, there is no cure for T1D and the main treatment consists of palliative life-long daily injections of exogenous insulin to maintain suitable blood glucose levels. This therapy is tedious to optimize to one's needs and is often associated with diverse complications, both acute (such as hypoglycemic and ketoacidosis episodes) and chronic (macro- and micro-angiopathies, e.g. vasculopathy, retinopathy, nephropathy, neuropathies).

1.1 Epidemiology

Nearly half a billion individuals have diabetes worldwide and T1D accounts for 9-10% of these patients[5]. Based on some recent meta-analysis, the incidence of T1D is estimated to be 15 per 100,000 persons[6], and its prevalence 19.5-43.8% worldwide[7].

Although T1D can occur at any age, most cases arise in children and adolescents, peaking between 3-6 and 10-14 years of age[8,9]. T1D incidence in children varies geographically, Norway, Finland, and Sweden being the most impacted countries (Figure 1)[10]. Studies of T1D incidence in children aged 0 to 14 years between 1990 and 1999 in 103 countries led to the conclusion of a global increase in incidence of 2.8% per annum over that period[8]. Similarly, in Europe, the incidence rate has been estimated to grow by 3.4% every year from 1989 to 2013. If sustained, this trend suggests a possible doubling in incidence in 18 to 25 years[11]. The national incidence rates range from 7.7 to 60.9 per 100,000 persons in Macedonia and Finland, respectively. In parallel in the US, the prevalence of T1D in children and adolescents between 2001 and 2009 increased by 21.1%[12] and is projected to triple by 2050[13].

In youth, T1D similarly affects females and males but tends to have a slight male predominance after puberty[14].

T1D is a major public health issue and, together with T2D, it has been recognized as one of the main causes of premature death and disability by the World Health Organization, making it one of the four priorities targeted by world leaders in the 2011 Political Declaration on the Prevention and Control of Non-communicable Diseases.

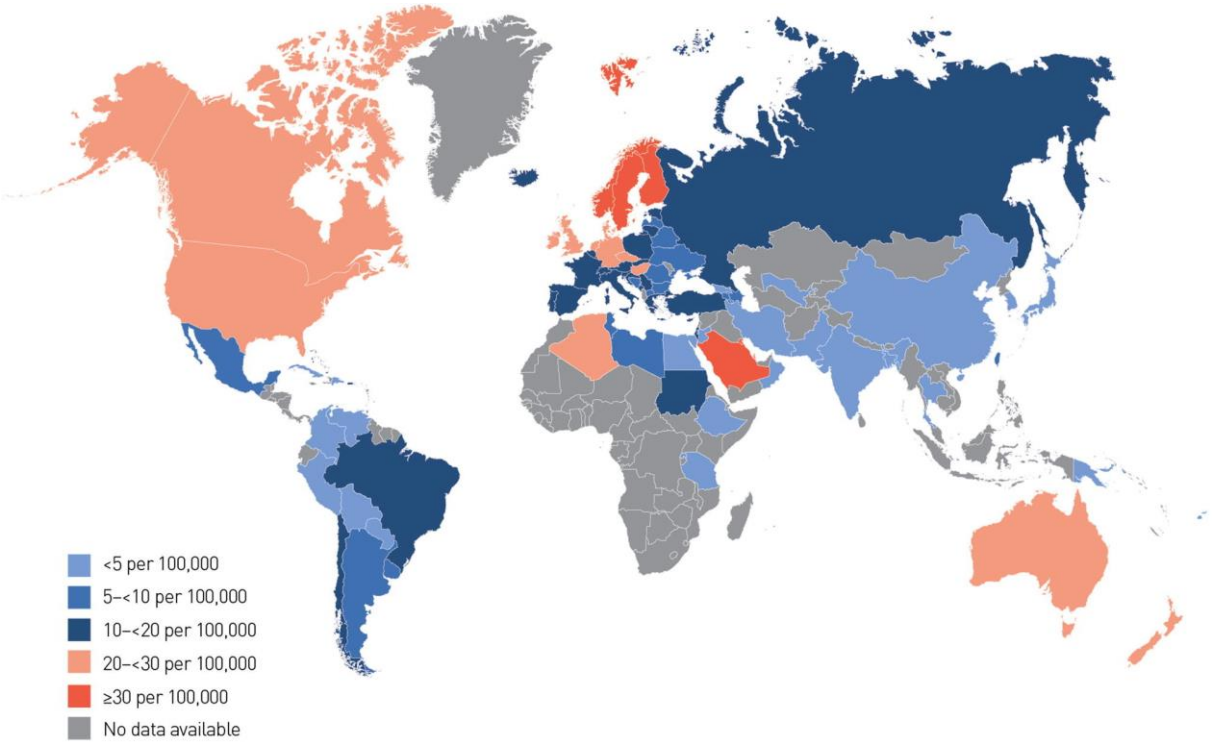


Figure 1. T1D epidemiology
Map of age-sex standardized T1D incidence rate (per 100,000) in children and adolescents (under 15 years old) (from [11]).

1.2 Natural history

The model of the natural history of T1D was first established in 1986[15] and has since been refined. This model suggests staged disease progression based on the loss of β-cell mass and function over time (Figure 2). Prior to the clinical onset, a long asymptomatic phase takes place, characterized by the gradual infiltration of islets by immune cells (insulitis) along with the production of autoantibodies (aAbs) and the activation of T cells recognizing β-cell antigens (stages 0 and 1). The initial β-cell destruction is triggered by these processes and dysglycemia can be manifested by an abnormal oral glucose tolerance test (OGTT) despite the absence of any symptoms (stage 2). As β-cell loss increases, hyperglycemia and associated symptoms (polyuria, polydipsia, weight loss) emerge, leading to clinical diagnosis (stage 3). Finally, only a few residual β cells remain in the pancreas, unable to support measurable insulin secretion, thus constituting the longstanding phase (stage 4).

As intelligible as this model is, it remains a simplified approach to the disease and does not represent its heterogeneity. First, it is widely accepted that genetic and environmental factors are at play in the progression from stage 0 to stage 1. These components are not elucidated yet, their existence relies on observations. Indeed, children raised in the bordering regions of two neighboring countries, with the same genetical background but different lifestyles, have a much higher incidence rate on one side of the border as compared to the other[16]. Moreover, people born in low T1D-incidence countries but migrating to high-incidence countries have a higher risk of developing the disease[17,18], which is more pronounced when migration takes place during childhood. Second, the development of T1D correlates with the nature and diversity of the aAbs, as the presence of 2 or more aAbs drastically increases the rate of clinical progression[19–21]. Third, the evolution towards clinical T1D depends on both the age at onset and the nature of the immune infiltrate. In children, T1D has been reported to advance faster, the immune infiltrate to be denser and the number of remaining insulin-containing islets (ICIs) to be lower as compared to adults[22–24]. Although the immune infiltrate and number of ICIs varies largely between individuals and across islets or pancreas sections of the same patient, these observations led to the definition of endotypes in T1D (see section 0). Finally, current imaging limitations prevent the accurate assessment of β-cell mass, which relies on an indirect measurement of β-cell function, although an observed viable β cell might not be functional.

1.3 Etiology

Over the last decades, evidence has been gained on the genetic, environmental, and metabolic components of T1D. A role for both genetical and environmental factors could be inferred early on upon the observation that the concordance in monozygotic twins is 30-65%[25,26]. On one hand, this advocates for a strong genetical determinant, but on the other hand supports a role for environmental events, as 35 to 70% of the risk remains unexplained despite an identical genetical background. Yet, T1D definite causes remain unknown and likely result from a combination of several factors. The current section focuses on the most probable triggers of the disease, for which consensus has been reached in the field.

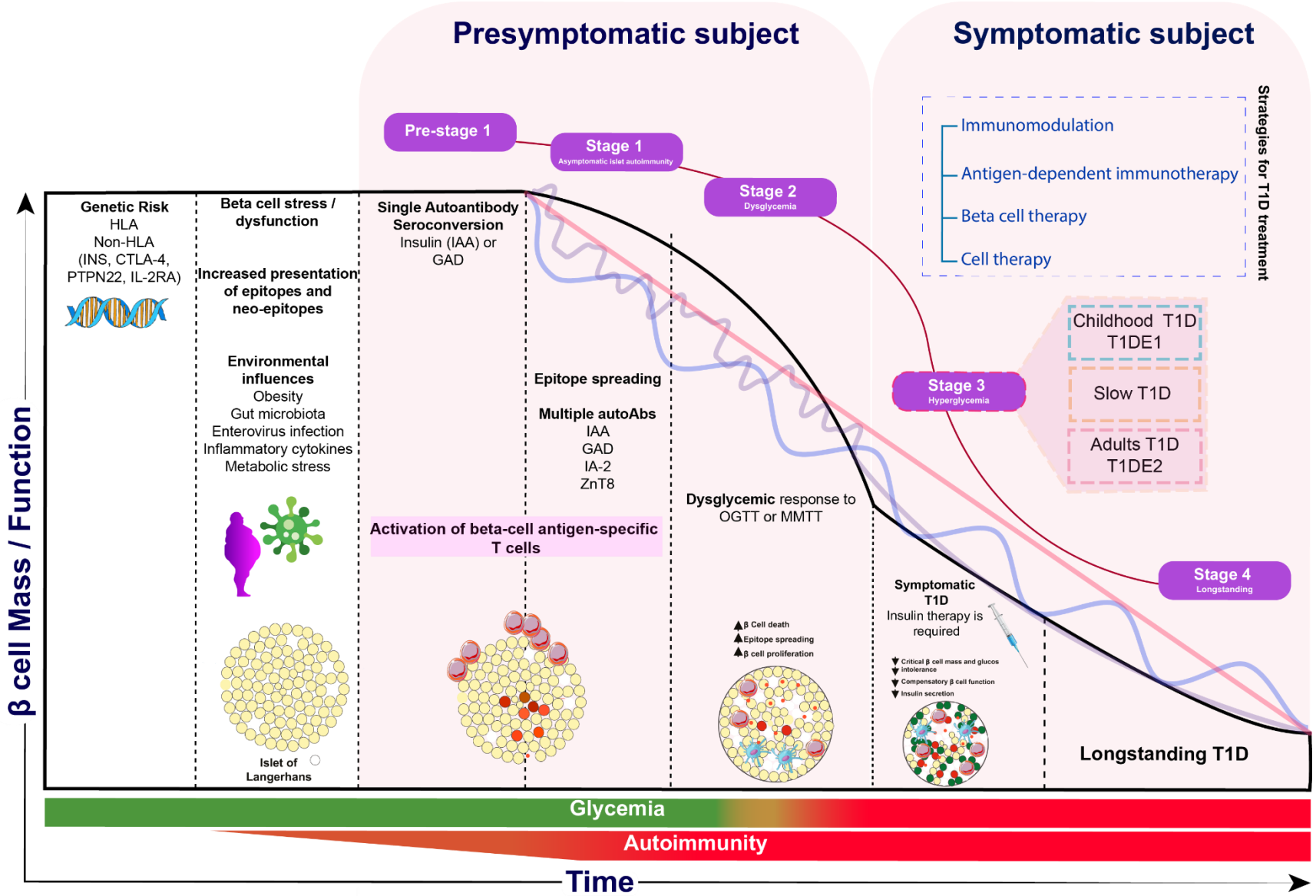


Figure 2. Natural history of T1D

Several genetic and environmental factors are at play in T1D etiology. The disease onset is preceded by a long asymptomatic phase, characterized by a gradual loss of β-cell mass and progressive infiltration of the islets by immune cells.

Stage 1 corresponds to the initation of the immune cells activation and appearance of the first aAbs. **Stage 2** is defined by the first occurences of dysglycemia, abnormal response to oral glucose tolerance test and the presence of multiple aAbs. **Stage 3** indicates the clinical diagnosis of T1D uncovered by hyperglycemia and metabolic imbalance. Exogenous insulin is required. Stage 4 describes long-standing T1D, with small residual β-cell mass.

1.3.1 Genetical factors

1.3.1.1 HLA genes

T1D development is strongly associated with the presence of certain alleles of the human leukocyte antigen (HLA) class II loci (Figure 3). More specifically, the DR3 (DRB1*03:01) and DR4 (DRB1*04:01) are widely accepted as susceptibility haplotypes and are most commonly inherited in linkage disequilibrium with the DQ2 (DQB1*02:01) and the DQ8 (DQB1*03:02) alleles, respectively. In addition, unaffected siblings sharing both DR3/DQ2 and DR4/DQ8 haplotypes with their diabetic sibling have an increased risk of 85% of developing the disease by the age of 15 as compared with siblings who do not carry both haplotypes[27]. On the contrary, the DR2/DQ6 (DRB1*15:01/DQA1*01:02-DQB1*06:02) haplotype is considered protective[28]. To a lesser extent, HLA Class I (HLA-I) alleles are also associated with T1D. The HLA-A*24:02 and -B*39:06 alleles are susceptible alleles whereas HLA-B*57:01 is a protective one[28,29]. However, only a minority of the total population carrying the susceptible HLA alleles will develop the disease. Overall, genes in the HLA region account for about 50% of the genetic weight in T1D, implying that other genes must be involved in T1D susceptibility[30].

1.3.1.2 Other genes

Most of the non-HLA loci predisposing to T1D relate to β-cell and immune functions (Figure 3). Expectedly, polymorphism in the insulin gene (INS) and more precisely in the variable number of tandem repeats (VNTR) in the promoter region correlates with variable risk of developing the disease. In the thymus, alleles constituted of short VNTRs lead to a lower transcription of INS mRNA and are associated with increased susceptibility to T1D. Reciprocally, alleles with long VNTRs induce high thymic INS mRNA synthesis, which in turn has a protective effect[31]. Both of these outcomes require homozygosity of the alleles and likely modulate the selection of insulin-reactive T cells and subsequent immune tolerance[32].

Additionally, genome-wide association studies (GWAS) have highlighted other genes outside of the HLA region associated with T1D, but the disease-causing variants remain largely unknown. Most of these genes are regulators of the immune response, such as PTPN22, IL2RA, and CTLA4[28,33] (Figure 3). Despite the predominance of immune-related genes, T1D is also a disease of the β cell. Indeed, up to 40% of T1D risk genes are expressed in human islets, where they participate in the β-cell response to the immune system[34]. In a human β-cell line, some of

these loci were also shown to be differentially expressed at the mRNA level upon exposure to inflammatory cytokines[35].

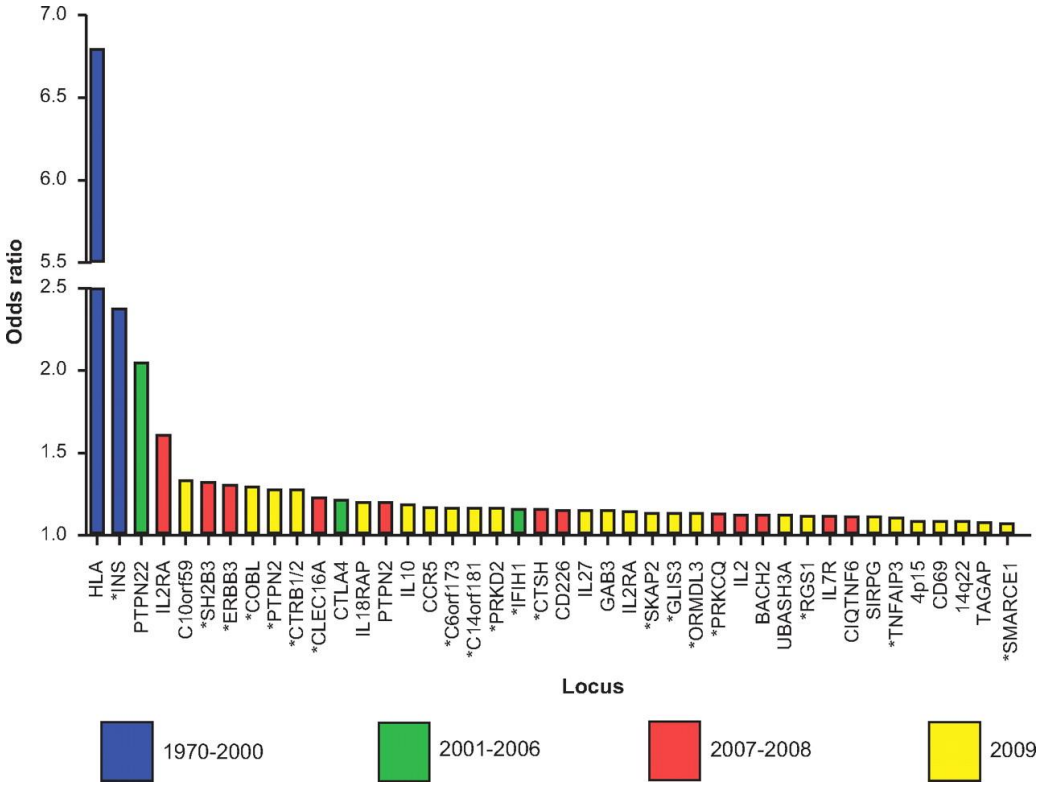


Figure 3. Genes associated with T1D susceptibility
The color code indicates the year of discovery and the asterisks (*) correspond to non-HLA genes expressed in human pancreatic islets (from [36]).

1.3.2 Environmental factors

The search for environmental triggers of T1D is an ongoing challenge tackled in particular by the TEDDY study (The Environmental Determinants of Diabetes in the Young, https://teddy.epi.usf.edu)[37]. Many different factors have been considered, from the prenatal environment to viruses, microbiota, and dietary elements[14,38–40] but the causal relationship between the disease and such determinants has, to date, never been proven. One of the difficulties lies in the timing of T1D diagnosis, which occurs too late as compared to the exposure to these possible triggers, making their detection challenging. Based on disease heterogeneity, it is unlikely that a single universal environmental trigger exists. Yet identifying the main environmental causes to develop prevention strategies becomes crucial, as the reported 3-4% of increase in the annual incidence rate cannot be explained by genetic drifting and suggests a heightened environmental pressure. This is further supported by studies showing that the weight of neutral and protective HLA alleles is declining[41].

1.3.2.1 Enteroviruses

The strongest evidence for an association with T1D onset implicates Enterovirus infection and more specifically Coxsackievirus B (CVBs)Annex 1[42–44] (reference 43 can be found in Annex 1). CVBs are non-enveloped single-stranded RNA viruses, known to have a tropism for the pancreas[45] and spread mainly through the oro-fecal route and occasionally through oral-oral or respiratory routes. They are common human pathogens causing, most of the time, flu-like or gastroenteritis symptoms. Nevertheless, in some cases, CVB infection can lead to serious diseases such as myocarditis, aseptic meningitis/encephalitis, and pancreatitis[46].

Based on the detection of viral RNA in the stools or neutralizing antibodies against CVB strains, several prospective cohorts such as the Finnish T1D Prediction and Prevention (DIPP) (https://dipp.fi) and TEDDY have reported a temporal association between CVB exposure and seroconversion and T1D progression[47–49]. On one hand, CVB1 infection is more prevalent in at-risk individuals before the appearance of aAbs, and on the other hand, prolonged CVB4 exposure is associated with a higher risk of seroconversion. Consistently, Enterovirus viral protein (VP)1 was found to have a higher frequency in islet sections of living donors with recent-onset T1D, from the Norwegian Diabetes Virus Detection study (DiViD)[50] and in tissue specimens from other cohorts[51,52], as compared to control subjects.

All of these studies suggest a causal relationship between CVB infection and T1D and one can envision several scenarios by which CVB could promote β-cell autoimmunity: 1) CVB is a lytic virus and its replication might kill the infected β cells, thus leading to the release of autoantigens. 2) Infected β cells could present viral peptides at their surface, in the frame of their HLA-I molecules, possibly activating anti-CVB cytotoxic $CD8^+$ T cells, similarly leading to β-cell death and autoantigens release. The flooding of the local microenvironment with autoantigens can trigger the bystander activation of T cells and further spark autoimmune processes. 3) A molecular mimicry between the sequences of a viral and a self-epitope could lead to the death of β cells by anti-CVB T cells. Such homologies have been previously described between CVB protein and the 65kDa form of glutamic acid decarboxylase (GAD65, also known as glutamate decarboxylase 2 or GAD)[53].

The clinical importance of CVB in T1D onset is further supported by recently launched clinical trials using CVB vaccination as a mean to possibly prevent T1D[54,55].

1.3.2.2 Microbiota

The association between gut microbiota and T1D originates from the observation that non-obese diabetic (NOD) mice (known to spontaneously develop diabetes and share genetic and immunological features with the human disease) have a higher chance of developing the disease if maintained in a pathogen-free environment[56,57]. Moreover, a reduced microbial diversity was found in children with islet autoimmunity, before progression to T1D[58,59]. Changes in the gut microbiota can affect gut permeability, leading to a so-called "leaky gut" thus allowing bacterial invasion in the tissue. The resulting inflammation could contribute to the activation of autoreactive T cells[60,61]. Also in this setting, the possibility of molecular mimicry mechanisms has been proposed[62].

1.3.2.3 Hygiene hypothesis

Current knowledge on the role of viruses and gut microbiota in T1D onset, together with the observation that industrialized countries with good sanitary conditions suffer from the highest incidence rates (Figure 1) has led to the hygiene hypothesis. This hypothesis speculates that limited exposure to microbes in early childhood could lead to autoimmunity. In other words, children exposed to a higher number and diversity of pathogens are more likely to build a strong immune response, preventing later detrimental effects[63].

1.4 Immunology of T1D

Despite T1D being clinically diagnosed as hyperglycemia rises, the long asymptomatic phase that precedes is characterized by the progressive invasion of islets by various immune cell types thus causing inflammation.

1.4.1 Islets immune infiltration

The abnormal presence of immune cells within the islets of T1D patients has been recognized decades ago[64]. It has since been extensively studied, both in human sections from cadaveric T1D donors and in the NOD mouse model. The former type of sample provides the most reliable observations, while the latter facilitates the characterization of the dynamics and chronology of the infiltration. Together with the recent technological advances such as imaging mass cytometry (IMC), which allows in-depth analyses of tissues[65], the nature and role of infiltrating immune cells have been better characterized along with the chronology of invasion.

Under physiological conditions, macrophages are found residing within the pancreatic islets, where they serve as scavengers, ensure primary immune surveillance and support tissue

homeostasis. Based on single-cell RNA-sequencing, a recent study has uncovered 5 subsets of macrophages present in the islets of NOD mice[66]. In these animals, the progression to diabetes correlated to the concomitant disappearance of the macrophages specialized in efferocytosis (clearance of apoptotic cells), a key process in the inflammation resolution[67] and the increase of cytokine-releasing pro-inflammatory macrophages harboring an M1 phenotype[66]. Macrophages are observed in the pancreata of both T1D patients and NOD mice in the early and late stages of the disease. They also appear to be the predominant infiltrating immune cells in insulin-deficient islets (IDIs)[66,68].

As they are not residing in the islets, the neutrophils can be considered the first immune cells to be recruited onsite. Indeed, they are found in the early stages of insulitis in NOD mice as well as in the pancreas of aAb-positive individuals[69]. They are the dominating immune cell type in longstanding T1D[65] (Figure 4). Their pathological role is indirect and resides in their capacity to form neutrophil extracellular traps (NETs)[69]. It has also been shown, in NOD mice, that neutrophils can release DNA-binding cathelicidin-related antimicrobial peptide (CRAMP). These peptides bind self-DNA, which results in plasmacytoid dendritic cells (pDCs) activation through toll-like receptor 9-myeloid differentiation factor 88 pathways and interferon (IFN)-α production in pancreatic islets[70]. The importance of macrophages and neutrophils has been further demonstrated by their depletion in NOD mice, which results in diabetes prevention[71]. The depletion of neutrophils only leads to the same outcome[70].

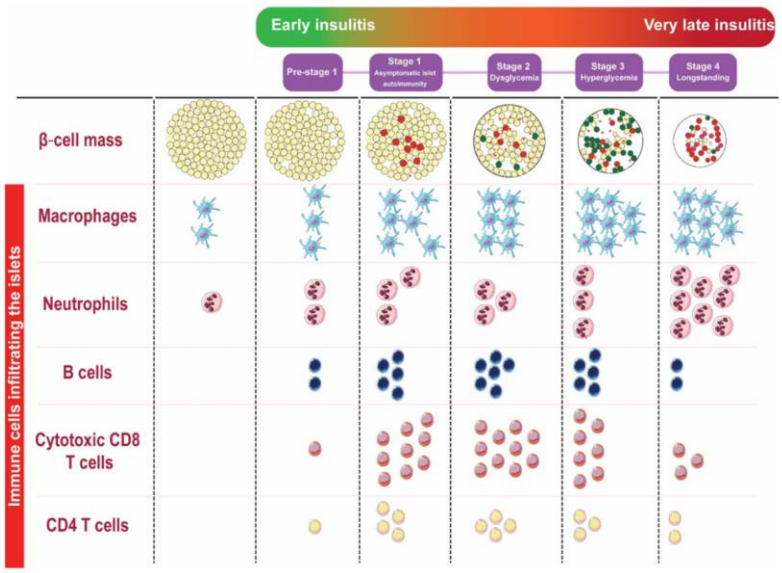


Figure 4. Insulitis evolution in T1D
Schematic representation of the immune cell types and their relative abundance within the pancreatic islets as β-cell mass decreases and paralleled to T1D staging.

The role of dendritic cells in T1D onset remains elusive[72], despite being found within the islets and in the exocrine tissue of the pancreas along with T cells[73,74].

B cells are also found infiltrating the islets. Their number increases as the disease progresses and β cells die (Figure 4). B cells have a crucial role in T1D, as they can uptake exogenous antigens through their surface IgGs and present them within their HLA-II molecules[75]. Moreover, NOD mice in which B cells have been depleted or invalidated for surface IgG expression are fully protected from T1D onset[76].

The most abundant cell type in the islet infiltrate are T cells (Figure 4). Recent IMC studies have highlighted the simultaneous recruitment of both $CD4^+$ T cells and $CD8^+$ T cells in β-cell-rich islets, pointing to the formation of a destructive immune pattern[65,77]. The abundance of cytotoxic $CD8^+$ T cells (CTLs) is steady across the disease stages but drops as the loss of β cells increases, and so do B cells and $CD4^+$ T helper cells (Figure 4). The observation that β2-microglobulin-deficient NOD mice (thus lacking HLA-I expression) do not develop autoimmune diabetes nor insulitis has underlined the role of CTLs in T1D development[78]. Moreover, the expression of cytotoxic effector markers such as granzyme B (GrB), IFN-γ, and CD107a by CTLs found in the islets of diabetes-prone NOD mice, is higher than the one of CTLs found in the periphery[79]. Also, a central role for islet antigen-specific $CD4^+$ T cells has been emphasized by mouse studies demonstrating disease prevention upon $CD4^+$ T-cell depletion[80]. Furthermore, islet-reactive $CD4^+$ T-cell clones can migrate to the pancreas of NOD mice and lead to the acceleration of diabetes onset[81]. Surprisingly, T_{regs} are nearly absent in the pancreatic sections of T1D and at-risk subjects[68,82] while they are detected early in the pancreatic immune infiltrate of NOD mice[83] and in a case report of slow-progressing T1D[84].

In summary, the immune infiltrate in T1D patients is characterized by a predominance of $CD8^+$ T cells followed by B cells and $CD4^+$ T cells and the presence of a high number of innate immune cells, mainly macrophages. But the overall heterogeneity observed across pancreatic sections of different donors has prompted the definition of guidelines for the diagnosis of insulitis[85]. The concomitant presence of a predominantly lymphocytic infiltrate defined with at least 15 $CD45^+$ cells/islet in at least 3 islets and the presence of pseudoatrophic insulin-negative islets are the two main criteria retained for this definition.

1.4.2 Cytokines

The variety of immune cells infiltrating the pancreas can play a crucial role in the development of T1D, as the cytokines they produce will shape the immune response.

Chronologically, the first cytokine signature observed even before the onset of T1D in at-risk children is dominated by IFN-α[86]. This is consistent with the increased expression of IFN-α also found in the pancreas of T1D patients close to the diagnosis[87]. Animal studies have confirmed the presence of both elevated IFN-α levels and IFN-α-producing pDCs in the pancreatic lymph nodes (pLNs) of young NOD mice[88(p)]. Blockade of IFN-α receptor leads to a delay in diabetes onset, highlighting the precipitating role of IFN-α in T1D development[88,89].

One of the earliest investigations tackling the role of cytokines in T1D showed the importance of IFN-γ and IL-6 in the late stage of the disease. Indeed, mice treated with anti-IFN-γ or anti-IL-6 antibodies had a significantly decreased diabetes incidence[90]. The role of IFN-γ was further emphasized by the decrease in β-cell injury in IFN-γ-deficient mice[91]. Nevertheless, the transfer of splenocytes isolated from these mice to NOD recipients passed on diabetes equally well as compared to splenocytes from diabetic NOD mice, suggesting that IFN-γ is not involved in the direct islet destruction but rather in the homing of diabetogenic T cells in the pancreas[92]. Alternatively, the synergism between IFN-γ and TNF-α has been suggested to be the final molecular effector responsible for β-cell apoptosis[93]. In line with this, a recent clinical trial with an anti-TNF-α antibody led to better endogenous insulin preservation and a reduced need for exogenous insulin[94].

While the above-mentioned cytokines are all produced by Th1 $CD4^+$ T cells, known to activate macrophages and drive cell-mediated immunity, the ones secreted by Th2 cells, responsible for strong antibody production and inhibition of several macrophage functions, have also been investigated. Early studies suggested that IL-10 cannot prevent autoimmune destruction of pancreatic tissue[95] but instead, its local production could accelerate the disease onset[96,97]. In contrast, its therapeutic use leads to the preservation of β-cell mass and function in NOD mice[98,99]. Moreover, low levels of IL-4 are found in both mice and recent-onset T1D patients[100,101]. But IL-4 has no role in diabetes aggressiveness as its complete elimination did not alter insulitis in NOD mice[102(p4)].

Despite many reports on the cytokines involved in T1D pathogenesis, the mechanisms remain poorly understood and their role is often controversial[103], possibly reflecting their redundant or context-specific functions to a large extent.

1.4.3 Humoral immune response: autoantibodies

Currently, the main biomarker indicating an autoimmune response against β cells is the presence of circulating aAbs directed against β-cell antigens. Their measurement is routinely used as a diagnosis tool[104]. Indeed, more than 90% of newly-diagnosed T1D patients are positive for one aAb while about 80% are positive for two or more[105].

The most described antigens targeted by aAbs are (pro-)insulin[106], GAD[107], the insulinoma antigen-2 (IA-2, also known as PTPRN or ICA512) which is a member of the protein tyrosine phosphatase (PTP) family[108] and the zinc transporter ZnT8 (SLC30A8)[109].

Of note, the aAbs usually appear years before clinical onset and seroconversion occurs early in life, with a median age below 2 years, supporting the initiation of β-cell autoimmunity long before T1D onset[110] (Figure 5). Many studies have investigated the use of aAb seroconversion as a prognosis tool, correlating the number of aAbs and age at seroconversion with the rate of progression to clinical T1D[3,111,112]. In genetically at-risk children, the progression to T1D at 10-year follow-up after aAb seroconversion is 70% in those with multiple aAbs and 14.5% in those with a single aAb while it is 0.4% in aAb-negative individuals[111]. These numbers increase even more when considering a longer follow-up (Figure 5), emphasizing the heterogeneity in the rate of disease progression.

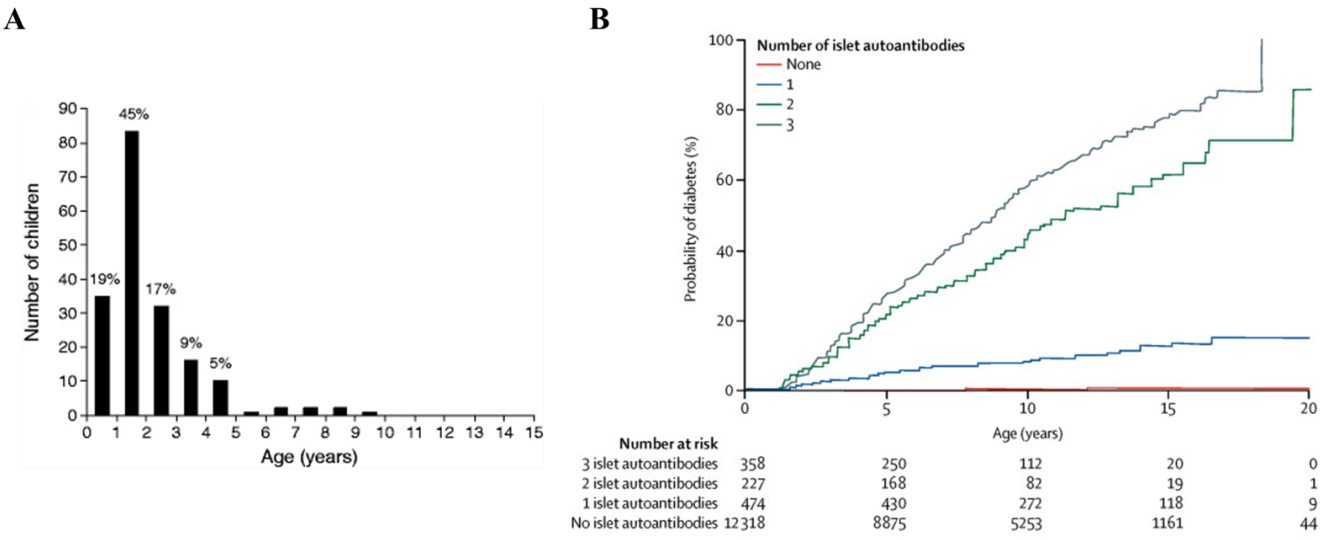


Figure 5. Autoantibodies in T1D
A. Age at seroconversion in children who progressed to clinical diabetes (adapted from [110]). **B.** The probability of developing T1D in children stratified by number of islet antibodies (adapted from [3]).

Despite being widely used as markers of disease progression, aAbs suffer from several limitations. First, their production does not play a notable role in T1D pathogenesis and instead

likely reflects the activation of aAb-secreting B cells, which is itself considered a secondary consequence of the self-reactive T-cell response. Second, a subtype of patients (up to 15%) harbors a clinical phenotype of T1D but remains aAb-negative, calling for additional diagnosis biomarkers[113]. Third, the timeframe during which multiple aAb-positive individuals are at high risk of developing T1D is too wide (10-20 years) to determine, with great accuracy, the best timing for some immune prevention strategies targeting stage 1 disease[114]. Fourth, only 10% (at 20 years) of single aAb-positive individuals progress to T1D (Figure 5). This population would highly benefit from improved risk stratification strategies to discriminate between progressors and non-progressors. Fifth, aAbs are classically used to select candidates for immune prevention or intervention trials but they are not useful monitoring tools for therapeutic efficacy since their titers change slowly, if at all[115].

1.4.4 Overview of T1D immunopathogenesis

All the evidence acquired advocates for the autoimmune nature of β-cell destruction in T1D: 1) specific HLA haplotypes are associated with the disease onset[30], 2) several aAb against β-cell antigens are found in T1D patients[105], 3) the pancreas is infiltrated by innate (neutrophils, macrophages, dendritic cells) and adaptive (T and B lymphocytes) cells, constituting the so-called insulitis[85], 4) the capacity of T cells to transfer diabetes in animal models[116] and 5) the encouraging efficacy of immunotherapies targeting T cells in delaying T1D[117,118].

The auto-immune process (Figure 6) starts in the thymus, which is the organ in which T cells selection and maturation occur. T cells recognizing self-antigens presented by thymic medullary epithelial cells are eliminated, thus limiting the odds of developing autoimmunity. However, this process of negative selection is marginal and several autoreactive T cells can escape into the circulation[119,120]. Alternatively, and depending on their affinity for the self-peptide, T cells avoiding apoptosis can differentiate into Forkhead box $P3^+$ ($Foxp3^+$) T cells, also known as regulatory T cells (T_{regs}). This cell type has a prominent role in the maintenance of peripheral self-tolerance and homeostasis as they control self-reactive T-cells that have escaped thymic selection[121].

In the pLNs, these self-reactive T cells might encounter antigen-presenting cells (APCs) migrating from the pancreas and harboring β-cell-derived peptides within their HLA molecules. The outcome of the binding of these T cells to their cognate peptide on APCs depends on the inflammatory conditions. In the absence of inflammation, the T cells remain quiescent or differentiate into T_{regs}. On the contrary, if innate immune cells are secreting cytokines or other

danger signals in the pancreas or in the pLNs, the self-reactive T cells are activated by their interaction with the APCs. Thus, the $CD4^+$ T cells promote the production of aAbs through their interaction with B cells. While $CD8^+$ T cells migrate to the pancreas, where they can recognize their target peptide at the surface of β cells, which suffer from an increased HLA expression favored by the inflammatory signals, resulting in their death. Finally, T_{regs} fail at eliminating self-reactive T cells, possibly due to a lack of suppressive activity and stability[122,123]

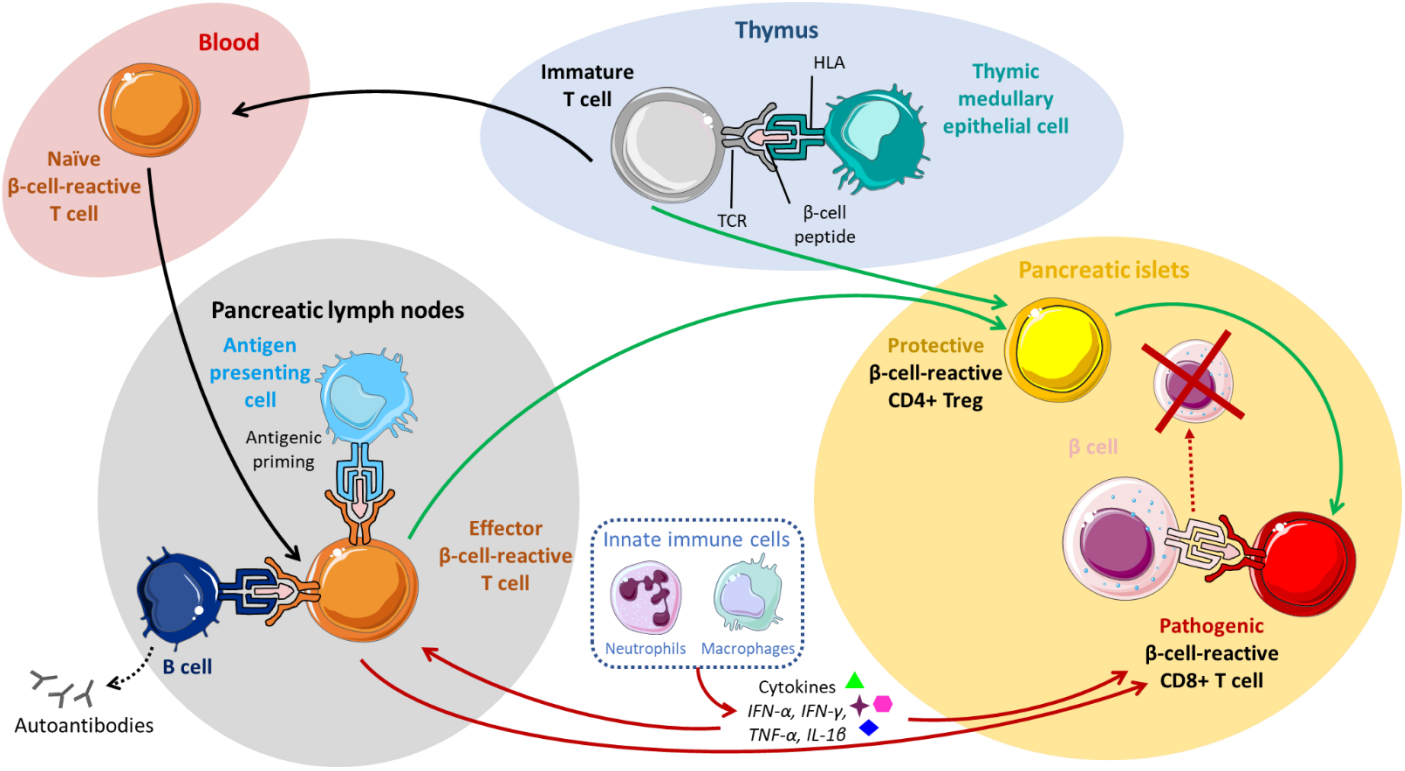


Figure 6. Schematic representation of T1D immunopathogenesis
In the thymus, a physiologically marginal clonal deletion leads to the escape in the circulation of self-reactive T cells. This first defective mechanism participates in T1D development. Once in the periphery, the self-reactive T cells can be activated as they recognize their cognate peptide presented by APCs. In inflammatory conditions, this interaction drives the pathogenicity of both $CD4^+$ and $CD8^+$ T cells, which promote the production of aAbs or participate in β-cell destruction in the pancreas, respectively.

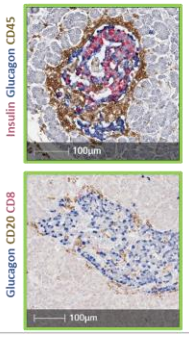
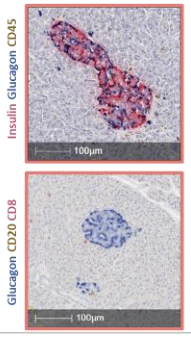
1.5 T1D endotypes

All the evidence gathered, regarding the residual β-cell mass, the nature of insulitis, the clinical history, and the variable therapeutic outcome of clinical trials highlights the heterogeneity of the disease and suggests the existence of "subtypes" of T1D. Histological studies on pancreas specimens from the Exeter Archival Diabetes Biobank (EADB) but also from the network for Pancreatic Organ Donors with Diabetes (nPOD) showed a correlation of the severity of insulitis with the age at onset, advocating for the definition of two endotypes: T1D endotype 1 and 2, T1DE1 and T1DE2, respectively[23].

As detailed in our recent review (Annex 2Annex 2)[24], T1DE1 defines individuals developing the disease at a young age (before 13 years). Histologically, their pancreata are characterized by few residual ICIs along with aberrant proinsulin processing and an infiltrate dominated by $CD20^{high}$ B cells and $CD8^+$ T cells[124]. They also tend to have a better response to immunotherapy[114,125–127]. On a genetic level, they frequently carry the HLA-DR4/DQ8 haplotype and first seroconvert for anti-insulin aAbs (IAA)[24]. On the other hand, T1DE2 is characterized by a later disease onset (>13 years of age), a higher number of residual ICIs along with normal insulin processing and reduced B-cell infiltrate ($CD20^{low}$), and milder $CD8^+$ T cell infiltration[24,128]. T1ED2 is frequently associated with the HLA-DR3/DQ2 haplotype and first seroconvert for anti-GAD aAbs[129]. The definition of these two endotypes further discriminates between a more aggressive (T1DE1) and a less aggressive (T1DE2) disease and also correlates with the two peaks of incidence reported, at 3-6 and 10-14 years of age[9]. All these features are summarized in Table 1.

Table 1. Characteristics of T1D endotypes

Adapted from [24]

	T1DE1	T1DE2
Age at onset	<13 years old	≥ 13 years old
Residual insulin containing islets (ICIs)	Few, infiltrated	Many, not infiltrated
Insulitis	>15 $CD45^+$ cells surrounding ICIs	Few 15 $CD45^+$ cells surrounding ICIs
CD20 expression in the pancreas	High	Low
Associated haplotype	HLA-DR4/DQ8	HLA-DR3/DQ2
First seroconversion	Insulin aAb	GAD aAb
Islet immunohistochemistry		

2 The HLA system

2.1 HLA class I and class II

2.1.1 Gene structure, expression, and function

The human leukocyte antigen (HLA) is the human version of the major histocompatibility complex (MHC). HLA genes are found on chromosome 6 and are highly polymorphic. They encode cell-surface proteins that regulate the adaptive immune system[130]. Three distinct regions have been identified in this gene (Figure 7): the classical HLA-I, encoding for the heavy chains of class I molecules, the HLA-II region, encoding for both α and β chains of the HLA-II molecules and the class III region encoding for components of the complements, TNF-α and others[130,131].

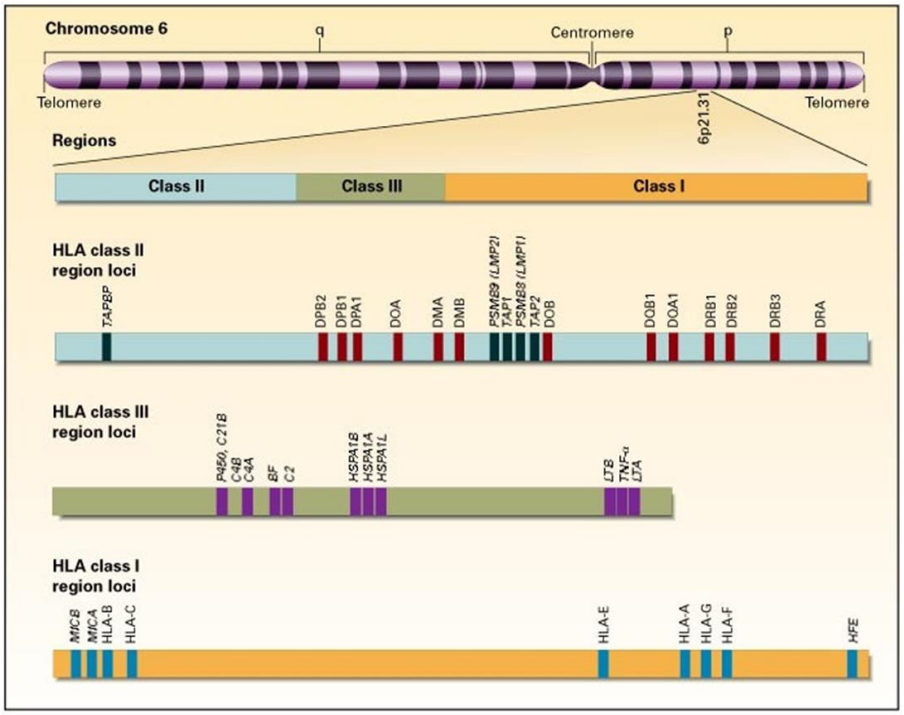


Figure 7. Location and organization of the HLA genes
The HLA complex is genetically divided into 3 regions: I, II, and III, each containing many loci, from which only the expressed ones are represented. Structurally or functionally, class III genes are not related to class I nor class II (adapted from [131]).

The class I molecules are found at the surface of nearly all somatic cells while the class II complexes are mainly expressed by APCs (B cells, dendritic cells, macrophages, and thymic epithelial cells)[131]. The function of both class I and class II molecules is to present pathogen-derived peptides to T cells, to initiate the adaptive immune response[131]. Their role is thus crucial for the discrimination between self and non-self. The polymorphism found in HLA molecules

is located at the peptide binding site[130]. Thus, each allele of HLA molecules binds a different repertoire of peptides, ensuring a wide variety of peptides presented at the cell surface.

2.1.2 Classical and non-classical HLA-I molecules

The HLA-I genes encode classical HLA-I, also referred to as Ia genes (HLA-A, -B, -C) and non-classical ones, also referred to as Ib genes (HLA-E, -F, -G).

The classical HLA-I molecules are glycosylated heterodimers, that typically bind short peptides (8-14 aa long) and are recognized by $CD8^+$ T cells. They are made of the heavy α chain (encoded on chromosome 6 as explained above) and the non-covalently bound light chain β2-microglobulin (β2m, encoded on chromosome 15).

The α chain consists of 3 domains, α1, α2, and α3 to which β2m is appended (Figure 8). The three-dimensional structure displays the peptide binding site at the top of the complex, located between two α helices from the α1 and α2 domains, resting onto a floor of a β-pleated sheet from the β2m chain[131,132]. The floor accommodates up to 6 pockets (A through F) distributed along the length of the groove, which serve as anchoring sites and are thus the main determinants of the peptide repertoire that a complex can bind[131,132]. Of note, the peptide size limit is determined by the presence of bulky aromatic residues at both ends, forcing the protrusion of some residues out of the groove and preventing the binding of long peptides[132].

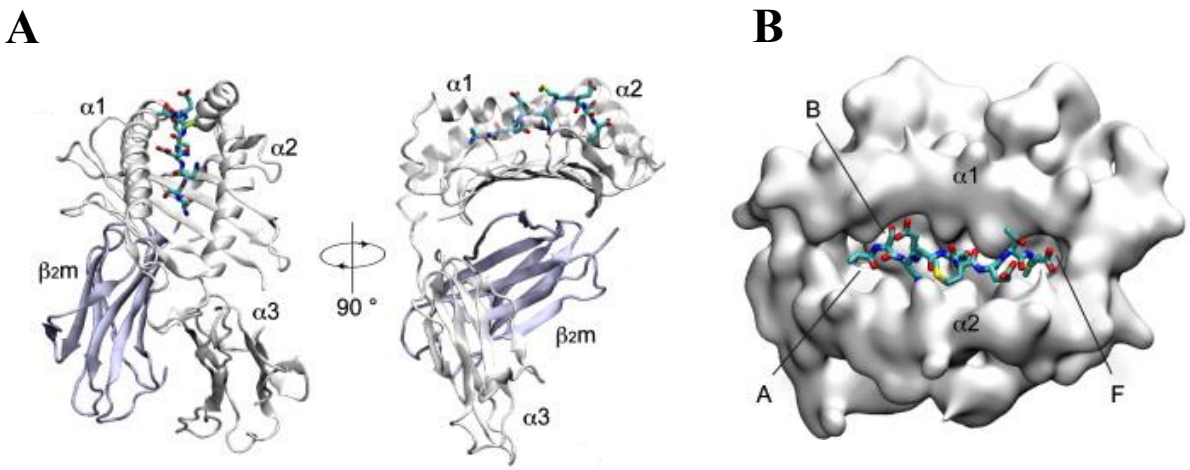


Figure 8. MHC-I structure
A. Mouse MHC-I (H-2D^b) structure in complex with the peptide ASNENMETM. The α chain is in grey, while the β chain is in blue and the peptide is pictured as a blue stick model. **B.** Modeling of the top view of the complex. The main pockets are shown, A, B, and F. (adapted from [132])

In T1D, HLA-I upregulation in ICIs has been a hallmark of the disease pathology, in comparison to surrounding acinar tissue and has been confirmed at both the RNA and protein

levels. This hyperexpression is observed irrespective of whether the islet is infiltrated by immune cells or not[133].

The non-classical HLA-I molecules are less polymorphic than their classical counterparts and present fewer isoforms. Their function can affect both the adaptative and the innate immune system. HLA-E is the least polymorphic HLA-I complex, it is structurally similar to HLA-Ia molecules and prefers short peptides (8-10 amino acids long). HLA-E binds to the leader sequences of HLA-I molecules (A, B, C, or G), which engages the interactions with CD94/NKG2A or CD94/NKG2C on NK cells, respectively inhibiting or activating NK cell-mediated killing[134,135]. But HLA-E can also be loaded with pathogen-derived epitopes, allowing its interaction with $CD8^+$ T cells[135,136]. It has also been found empty at the cell surface, yet receptive to peptides[137]. A soluble isoform of HLA-E exists, that provides protection for bystander cells and is induced by inflammation[138]. In T1D, HLA-E expression is upregulated upon pro-inflammatory mediators and is rather specific to ICIs[133,139]. Polymorphisms at the HLA-E locus are associated with age at T1D onset[140]. HLA-F molecules harbor a diverse repertoire of longer peptides (16 aa) and can also interact with NK cells[141]. In islets of T1D donors, HLA-F expression is upregulated at the mRNA and protein levels[133,139,142]. Several membrane-bound or soluble isoforms of HLA-G have been reported. Like other HLA-Ib molecules, HLA-G has an immune-modulatory role and can interact with $CD4^+$ T cells, $CD8^+$ T cells, NK cells, dendritic cells, and T_{regs}[143]. In T1D primary islets, HLA-G is found both at the cell surface and in a subset of insulin-containing granules[133,139].

2.1.3 Classical and non-classical HLA-II molecules

The HLA-II region encodes for classical HLA-II molecules (HLA-DP, - DQ, -DR) and for non-classical ones (HLA-DM, -DO). The classical HLA-II molecules are heterodimeric glycoproteins recognized by $CD4^+$ T cells. They consist of an α and a β chain, each containing 2 domains (α1, α2, and β1, β2) that form two immunoglobulin-like domains topped by a peptide-binding site (Figure 9)[144]. Quite similarly to HLA-I proteins, the peptide-binding groove is floored by a β-sheet and is laterally limited by two α helices, formed by the α1 and β1 chain[132]. The peptide holds tightly into the HLA-II molecule through its interaction with several pockets found at the bottom of the binding groove, the major ones being P1, P4, P6, P7, and P9[132]. Contrary to the HLA-I peptide binding site, which harbors closed extremities, the HLA-II groove is open at both ends. This conformation ensures the binding of a core sequence in the pockets, yet accommodates long peptides (usually 12-16 aa long, but up to 20)[132,144].

The non-classical HLA-II molecules are structurally close to the classical HLA-II but they differ in their function. Indeed, the non-classical HLA-II complexes are known to regulate peptide loading into classical HLA-II.

The expression of HLA-II molecules by target β cells in T1D has been strongly debated for decades. The most recent line of evidence confirms that β are indeed able of harboring HLA-II complexes, which is likely a direct consequence of the inflammatory environment[145,146]. But to date, the expression of co-stimulatory molecules, such as CD80 or CD86 could not be demonstrated suggesting that despite this HLA-II presentation, β cells cannot act as APCs in T1D.

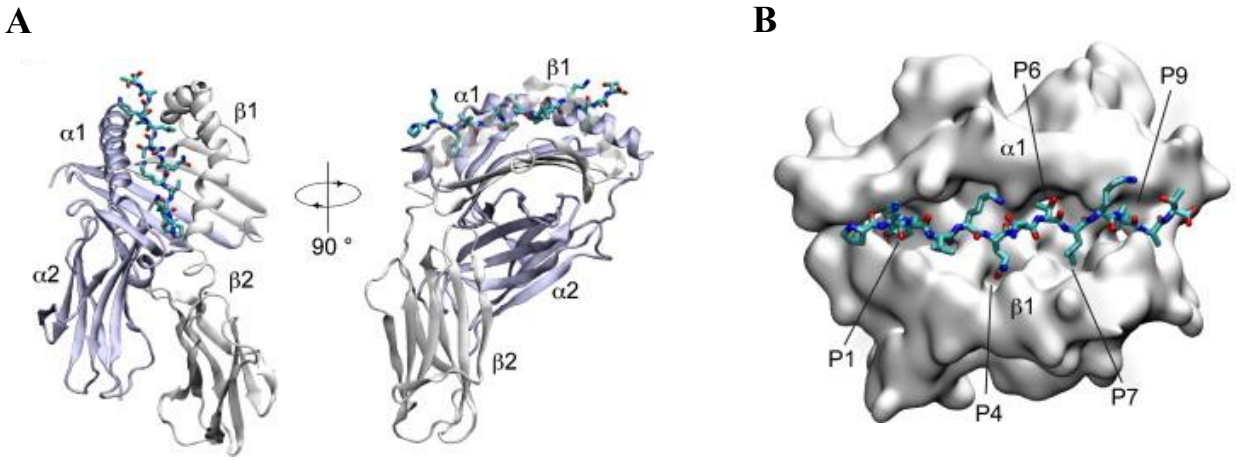


Figure 9. MHC-II structure
A. HLA-DR1 structure in complex with the peptide PKYVKQNTLKLAT. The α chain is in grey, while the β chain is in blue and the peptide is pictured as a blue stick model. **B.** Modelling of the top view of the complex. The main pockets are indicated, P1, P4, P6, P7, P9 (adapted from [132])

2.2 Antigen processing and presentation pathways

While the HLA-I antigen processing and presentation (APP) pathway revolves around peptides from intracellular cytosolic proteins, the HLA-II APP is mainly articulated around peptides from extracellular origin. Nonetheless, several intersections between these pathways exist. Peptides from extracellular origin can be presented onto HLA-I molecules through a process named cross-presentation[147,148] and peptides from intracellular origin can be loaded onto HLA-II complexes by autophagy[149]. This section aims at giving a global, yet non-comprehensive, view of these pathways.

2.2.1 HLA-I antigen processing and presentation

2.2.1.1 Classical pathway

In the cytosol, old proteins as well as defective ribosomal products (DRiPs), which can be misfolded or unfinished proteins, are poly-ubiquitinated allowing their degradation by the proteasome into smaller fragments[150]. The proteasomal degradation is key to the formation of a C-terminus that is compatible with HLA-I binding while the N-terminal stretch is of variable length, resulting in peptides of 3 to 18-22 aa long[151–153]. While the vast majority of these peptides undergo further degradation into single amino acids that can be re-used by the cell, a portion of them reach the transporter associated with antigen processing (TAP) and translocate from the cytosol to the endoplasmic reticulum (ER)[153–155] (Figure 10).

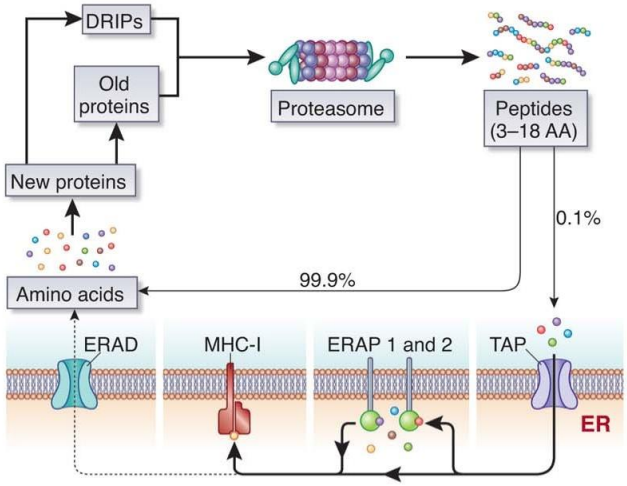


Figure 10. The peptide cycle for HLA-I

In parallel, the newly-translated HLA-I molecule first interacts with two chaperones, BiP (binding immunoglobulin protein, also known as GRP78) and calnexin, which start its folding and assembling into the ER. Later, the association of the β2m subunit is concomitant to the replacement of calnexin by calreticulin. The peptide-loading complex, which orchestrates the transport and loading of peptides onto the HLA-I molecule, is fully formed upon the binding to TAP, tapasin, and ERp57[154,155].

Once in the ER, the N-termini of the peptides are further trimmed by the ER aminopeptidase (ERAAP or ERAP1) and loaded onto the HLA-I molecule[156]. Then, the peptide-HLA-I (pHLA-I) complex dissociates from all the chaperones and is transported toward the cell surface *via* the secretory pathway[154,155]. The APP pathway of HLA-I peptides is represented in Figure 11.

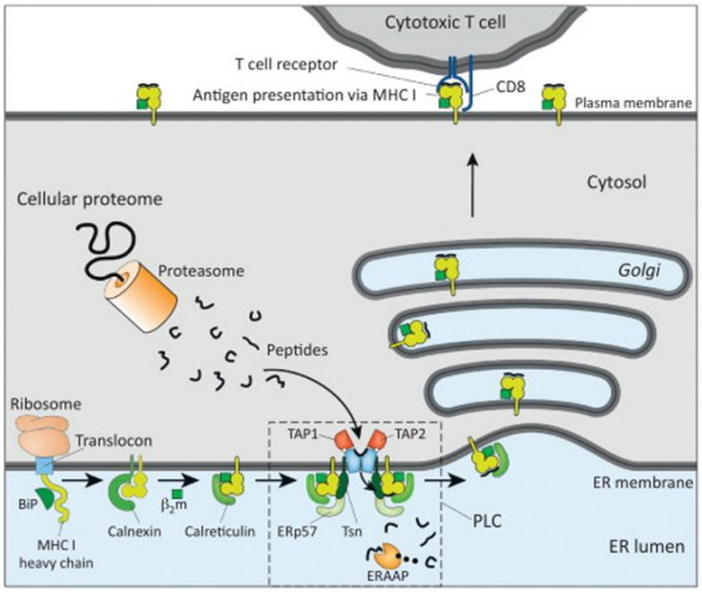


Figure 11. Conventional HLA-I antigen processing and presentation pathway
The transmembrane domain of TAP is colored in blue while its nucleotide-binding domain is in red. BiP: binding immunoglobulin protein; TAP: transporter associated with antigen processing; Tsn: tapasin; PLC: peptide-loading-complex (from [154])

2.2.1.2 Alternative antigen sources and presentation pathways

First, several studies have shown that HLA-I peptide presentation can occur in a proteasome- and/or TAP-independent manner[157–159]. Furthermore, other cytosolic proteases, such as insulin-degrading enzyme, tripeptidyl peptidase II and caspases, can participate in the generation of peptides for HLA-I presentation[160–162]. Some reports also suggest that peptides could be transferred from one cell to another through gap junctions[163,164]. But the contribution of these pathways to the overall HLA-I APP remains poorly evaluated and likely minimal. The main alternative antigen sources and presentation pathways reside in the induction of the immunoproteasome, cross-presentation mechanism, and vesicular pathway.

2.2.1.2.1 Role of the immunoproteasome

The classical 26S proteasome is a barrel-shaped protein complex constituted of a catalytic core, the 20S particle, capped by a regulatory 19S subunit on each side. Within the 20S core, the catalytic activity is ensured by the β1, β2, and β5 subunits which have a caspase-like (acidic residues), trypsin-like (basic residues), and chymotrypsin-like activity, respectively[165].

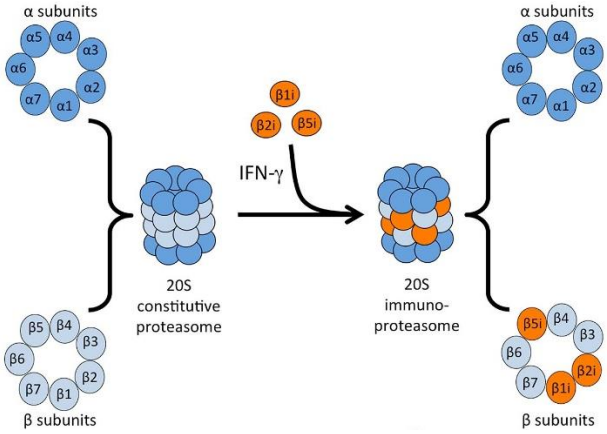


Figure 12. Formation of the immunoproteasome
The 20S catalytic core of the proteasome is made up of two outer α rings and two inner β rings. Upon exposure to inflammatory cytokines, such as IFN-γ, the three subunits β1i, β2i and β5i are synthesized and incorporated into newly formed proteasomes, replacing their constitutive counterparts and forming 20S immunoproteasomes (from[166]).

Upon inflammation, the three immunoproteasome subunits β1i (LMP2), β2i (MECL-1), and β5i (LMP7) are incorporated in place of their constitutive counterparts to form the immunoproteasome (Figure 12). The assembly rate of the immunoproteasome is four times faster than the one of the constitutive proteasomes, allowing a rapid response to inflammatory stimuli[167]. IFN-γ is the most potent inducer of the immunoproteasome expression, but IFN-α, IFN-β, and TNF-α can also lead to its expression[167]. Whether the constitutive proteasome and the immunoproteasome diverge in function is still debated. It was first suggested that both complexes work differently and generate peptides of different nature[168,169]. On the opposite, more recent reports imply that they use the same cleavage sites but have distinct substrate specificities[170–172]. Despite the lack of consensus regarding their cleavage properties, it is widely admitted that the induction of the immunoproteasome leads to increased proteolysis and overall changes in the type of peptides presented within HLA-I complexes.

2.2.1.2.2 Cross-presentation

The mechanism of cross-presentation consists in the presentation of exogenous peptides within the HLA-I molecules. This process has been extensively studied in APCs and results in the stimulation of naïve $CD8^+$ T cells into activated CTLs by cross-priming.

First, the APC internalizes extracellular antigens through phagocytosis or receptor-mediated endocytosis. These antigens are then directed either towards the vacuolar or the cytosolic pathway[173,174]. In the former, the phagocytic vesicle undergoes acidification and activation of lysosomal enzymes, leading to the degradation of the internalized antigens into peptides. The peptides can be docked into the HLA-I molecules within the phagosome and

exported to the cell surface (Figure 13). It remains unclear whether the HLA-I molecules originate from the ER or are recycled from the surface. In the cytosolic pathway, the newly-engulfed antigens are transported from the endocytic vesicle to the cytoplasm, by unknown mechanisms. Once in the cytosol, the antigens can be degraded by the proteasome and either engage in the classical APP pathway into the ER or be transported back into the endosome for HLA-I loading and presentation of the complex at the cell surface (Figure 13)[173,174].

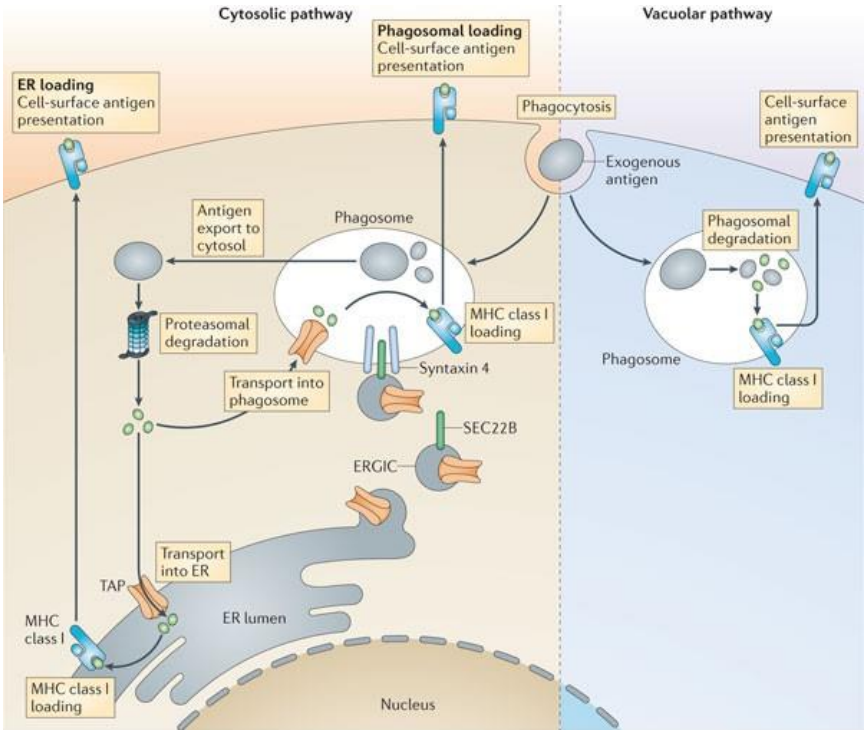


Figure 13. Antigens cross-presentation by dendritic cells
Exogenous antigens are phagocytosed and can be exported into the cytosol (cytosolic pathway), where they are degraded by the proteasome. The resulting peptides are loaded onto HLA-I molecules within the ER or after re-importation into the phagosome. If not exported into the cytosol, the phagocytosed antigens can be processed and loaded on HLA-I molecules directly in the phagosome (vacuolar pathway) (from [173]).

2.2.1.2.3 Vesicular pathways

Irrespective of the pathway through which pHLA complexes reach the cell surface, their transit systematically relies on the vesicular trafficking of the cells. Besides, as any cell surface receptor, the complexes are continuously recycled and remain presented at the cell surface only for a given duration[175], which depends on the peptide affinity and the overall stability of the complex[176]. Thus, constant exchange between the cell surface, the Golgi, and the ER, that exploits the vesicular pathway exist in every cell[175] and could constitute a hub for alternative APP.

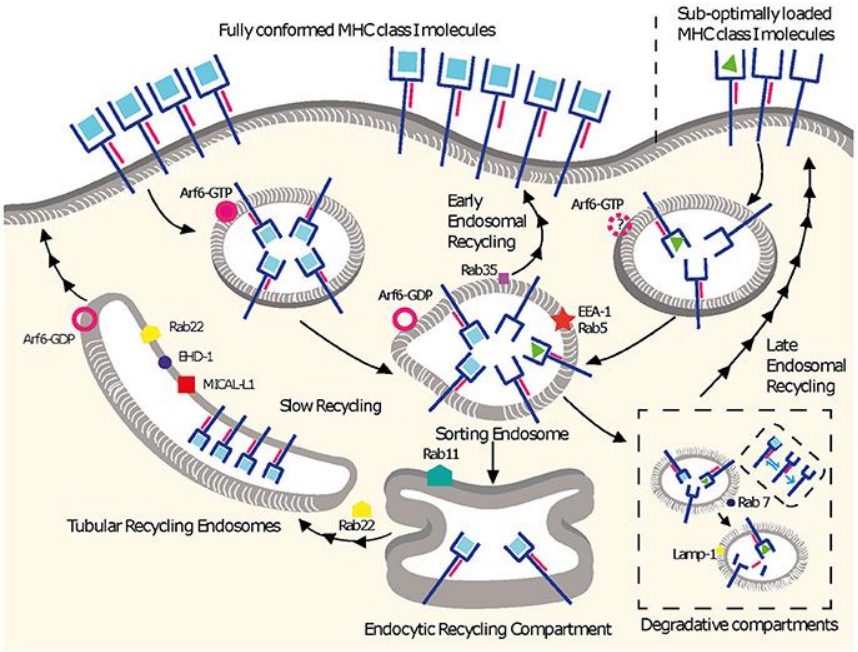


Figure 14. Recycling pathways in non-professional APCs
Depending on the affinity of the peptide they carry, the HLA-I molecules can be recycled through different pathways. Stable complexes transit to sorting endosomes, where they are further addressed to a slow or early recycling. Sub-optimally loaded complexes also arrive to sorting endosomes and are degraded or undergo late endosomal recycling (from [175]).

The most stable pHLA-I complexes can be recycled by clathrin-independent endocytosis[177] and are directed toward sorting endosomes, where they will be directed towards either slow or early endosomal recycling (Figure 14)[175]. Sub-optimally loaded complexes are also internalized and moved to sorting endosomes but the detailed pathway is not yet fully elucidated. Part of these complexes is degraded in dedicated compartments, while others can be recycled to the cell surface (Figure 14)[175].

2.2.1.2.4 In β cells

The implication of the alternative APP mentioned above in β cells and their possible role in T1D has been more specifically covered in our recent review (Annex 3)[178].

2.2.2 HLA-II antigen processing and presentation

2.2.2.1 Classical pathway

The presentation of HLA-II ligands at the cell surface requires the uptake of extracellular proteins into vesicular compartments for their degradation and the binding of the formed peptides to HLA-II molecules.

Three routes allow the internalization of extracellular antigens along with efficient sorting to the vesicular machinery. First, clathrin-mediated endocytosis promotes the uptake of

cell surface complexes, membrane-bound proteins, and soluble macromolecules. Second, extracellular particles (antigens, pathogens) are selectively engulfed by phagocytosis. Third, macropinocytosis allows the non-selective capture of large quantities of extracellular components[155]. The endocytic vesicles by which these antigens penetrate the cells then mature into late endosomes and lysosomes. This progression is characterized by luminal acidification and the fusion with enzyme-rich vesicles, which altogether actively stimulate antigen degradation and the formation of peptides. The pH regulates the activity of the proteases found in these heterogeneous vesicles, such as gamma-interferon-inducible lysosomal thiol reductase, and cathepsins[155].

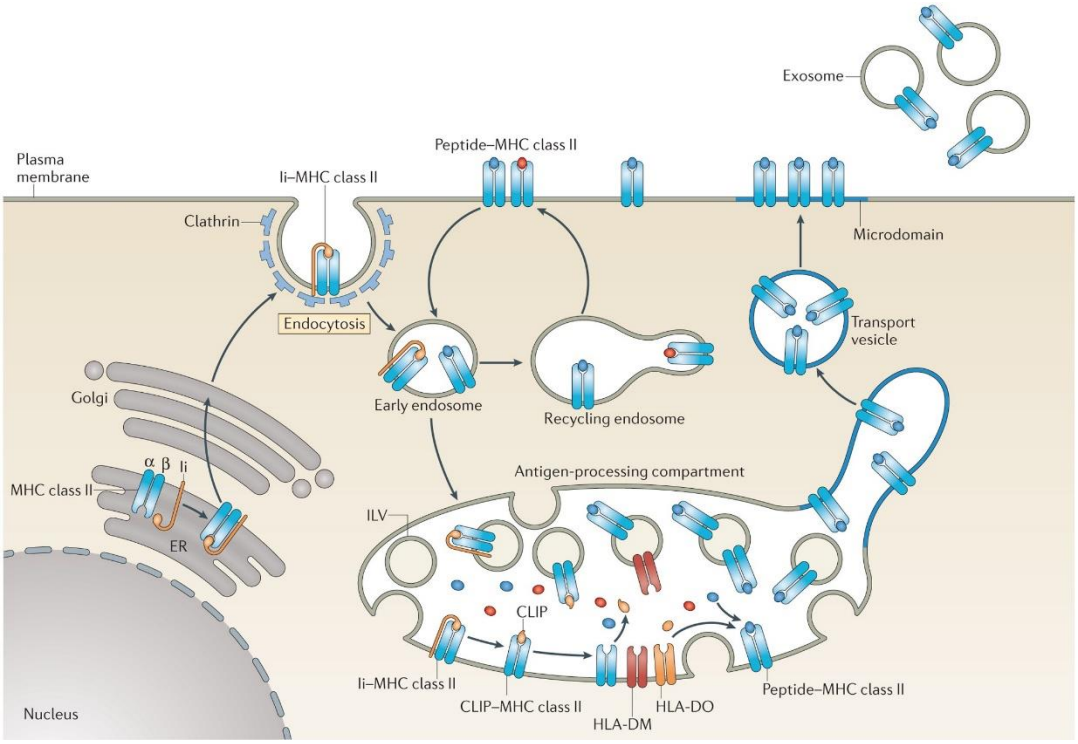


Figure 15. Conventional HLA-II antigen processing and presentation pathway
Ii: invariant chain; CLIP: class II-associated invariant chain peptide; ILV: intraluminal vesicle. From [179]

In parallel in the ER, the α and β chains of HLA-II molecules are concomitantly synthesized and associated with each other and with the invariant chain Ii (Figure 15)[155,180]. The Ii chain serves as a signal molecule to guide the newly-formed complex across the diverse vesicular compartments and protects its peptide-binding groove until it reaches the late endosome for peptide capture. The HLA-II complexes arriving in the late endosome can originate from i) direct trafficking from the trans-Golgi network after biogenesis, ii) endocytosis of newly-formed complexes that were exported at the plasma membrane, iii) a recycling endosome from the plasma membrane[155,179]. Once in the late endosomes, the Ii chain is released

by proteolysis, leaving only a 24-amino acid long class II-associated invariant chain peptide (CLIP) in the peptide-binding cleft[180]. The exchange of CLIP for a higher affinity ligand is promoted by the HLA-DM molecule, a protein that structurally resembles conventional HLA-II heterodimers (Figure 15). The activity of HLA-DM is modulated by another heterodimer, HLA-DO, but the detailed mechanism remains unknown[155,179]. In NOD mice, the overexpression of H2-DO (orthologous gene of HLA-DO) reduces T1D incidence, seemingly due to an impaired presentation of antigenic peptides within HLA-II molecules in the endosomes[181,182].

Finally, peptide-loaded HLA-II complexes are transported to the cell surface for the presentation of their peptide to $CD4^+$ T cells.

2.2.2.2 Alternative antigen sources and presentation pathways

The HLA-II APP pathway relies on endo-lysosomal trafficking, a system that serves several purposes. One of them is autophagy, an intracellular mechanism aiming at recycling self-proteins by lysosomal degradation. As exemplified by the cross-presentation of exogenous antigens onto HLA-I molecules, the many functions of the vesicular network can intersect with one another. Similarly, autophagy can feed the HLA-II APP pathway[183,184]. Approximately 20-30% of HLA-II-bound antigens are of cytoplasmic or nuclear origin in human B cells and DCs[185].

Other non-autophagic pathways can contribute to non-classical HLA-II peptide presentation. One of them involves major players of the classical HLA-I APP route, implying that the HLA-II hijacks the HLA-I APP pathway. TAP was shown to play a role in the HLA-II presentation of a short cytosolic viral peptide, suggesting that cytosolic peptides could be translocated to the ER in a TAP-dependent manner. Moreover, stress conditions can imbalance the ratio between HLA-II molecules and the protective Ii chain, resulting in the binding of partially or unfolded proteins to HLA-II molecules[186]. Some focused studies could further support a role for TAP, for the proteasome but also for HSP90, a protein involved in heat-shock protection[187].

2.3 The immunopeptidome

2.3.1 Definition and identification

The immunopeptidome can be defined as the repertoire of peptides presented within the frame of HLA-I and -II molecules at the surface of a cell. Its composition depends on the HLA allotypes and reflects the cellular status (stress, high metabolic demand, cytokine exposition, viral infection, etc)[188–190].

HLA molecules drive peptide diversity through their high polymorphism. Currently, there are three well-established ways to characterize the binding preferences of HLA molecules. First, sequences from publicly available databases, such as SYFPEITHI[191] and IEDB[192], can be used to test synthetic peptides in *in vitro* binding assays. Second, pHLA complexes can be eluted from cell surfaces and peptides can be sequenced by liquid chromatography and tandem mass spectrometry (LC-MS/MS), a technology named immunopeptidomics[193]. Third, the sequence of known binders together with the structural properties of the binding groove and non-covalent interaction between residues can be bioinformatically computed to develop *in silico* prediction algorithms[194].

2.3.1.1 Binding prediction algorithms

Each HLA allele has a distinct peptide binding specificity at the amino acid level, from which motifs arise. These motifs factor in the bond energy of amino acids at each position along with the influence of the neighboring residues and, as a result, reveal specific amino acids that anchor the antigen in the binding groove[195,196].

Using artificial neural networks, these parameters have been incorporated into several prediction algorithms, such as NetMHCpan, NetMHCpanII[194], and MHCflurry[197], which are constantly being fed and improved with new datasets made available. Nonetheless, a myriad of novel binding prediction tools keeps emerging (mixMHCpred[198], mixMHC2pred[199], NNAlign-MA[200], NetCleave[201]), each having their specificities, emphasizing the need for a consensual and reliable algorithm.

Due to the higher flexibility of the HLA-II binding groove, HLA-II prediction algorithms are more complex to develop and thus less established than HLA-I ones, despite recent advances[199,200,202]. As most of the data available to train these algorithms come from cancer datasets, the resulting predictions are prone to carry some intrinsic biases and to be less effective in other disease settings, such as autoimmunity. It should also be underlined that an

antigen predicted to be a good binder, does not make it a good epitope (i.e. well recognized by T cells) whereas a good epitope usually displays high-scoring predictions. Despite these limitations, prediction algorithms remain widely used to complement mass-spectrometry (MS)-based strategies.

2.3.1.2 Immunopeptidomics

2.3.1.2.1 Workflow

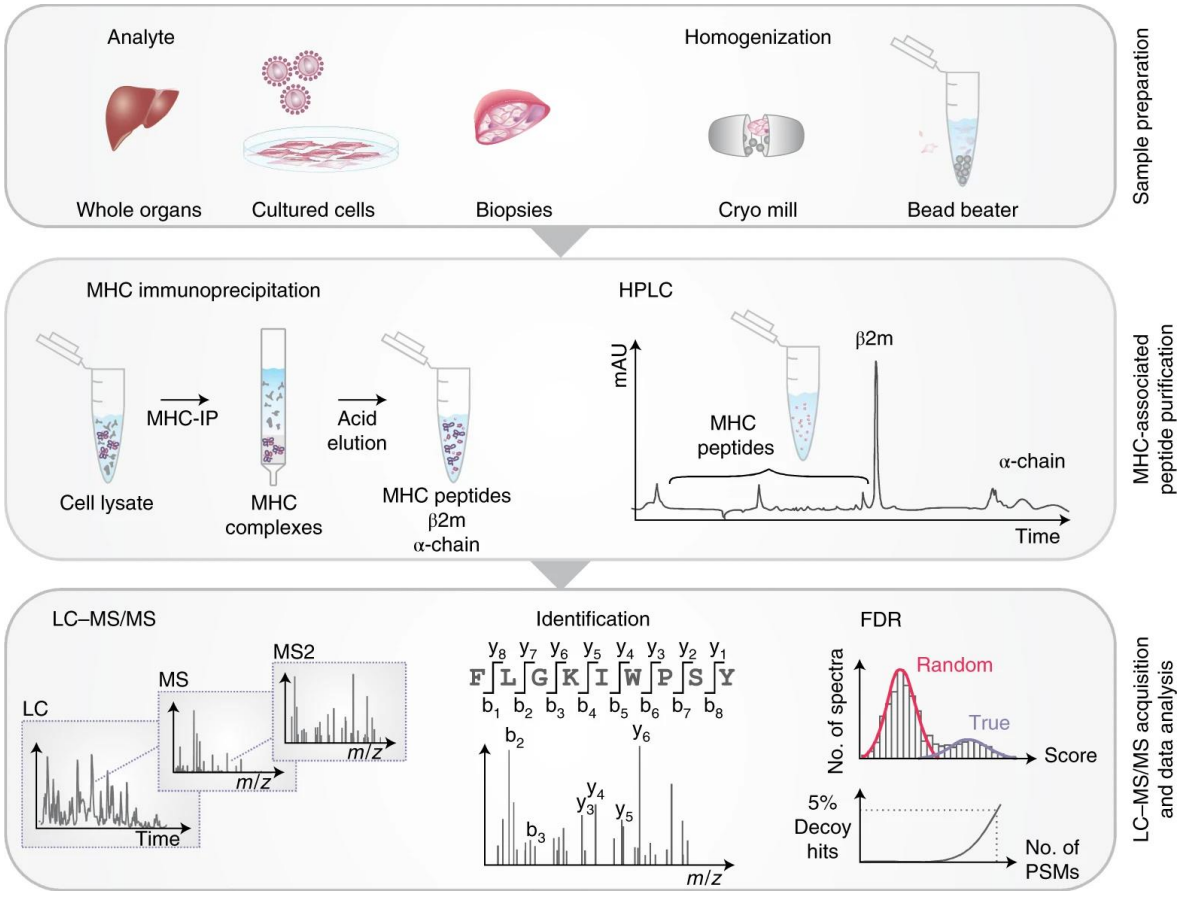


Figure 16. Workflow for the purification and identification of HLA-bound peptides by LC-MS/MS
Adapted from [193]

The only strategy for the identification of HLA ligands, that allows a complete view of the antigens truly presented at the cell surface is immunopeptidomics (Figure 16). This technique requires a considerable amount of biological material to be lysed. pHLA complexes are then eluted, either by mild acidic elution or by immunoprecipitation (IP) using antibodies specific for the HLA molecule[193]. The IP is the preferred option as it diminishes the risk of contaminations and increases the peptide yield[203]. The peptides are eluted from the complexes in an acidic environment and purified, before being separated by nano-LC and analyzed by MS/MS[193]. Immunopeptidomics experiments require the use of a high-mass-accuracy analyzer

to ensure reliable peptide identifications and a mass spectrometer configured for this specific use, which differs substantially from classical proteomics[193,204]. Finally, peptide identification is performed by dedicated softwares and is further detailed in section 2.3.1.2.3.

With this approach, the most abundant peptides are the easiest to identify and might be overrepresented at the expense of rare sequences, such as neo-antigens[205]. Currently, the only way to compensate for this limitation is to use a large amount of starting material.

2.3.1.2.2 Type of acquisitions

There are two major acquisition types in MS: data-dependent acquisition (DDA) and data-independent acquisition (DIA).

In DDA, during each cycle, the instrument selects the top N (usually 10 or 20) most intense ions and individually fragments and analyzes them[206]. This unbiased approach is commonly used for discovery experiments. This strategy suffers from limited reproducibility across technical replicates (35-60%)[207].

On the contrary, DIA suits hypothesis-driven experiments and allows the detection and quantification of low abundant species. In that case, the instrument does not proceed with any prior selection and fragments all ions in each cycle[206]. The reproducibility is better than DDA but the resulting spectra are multiplexed and require more complex data analysis[208].

2.3.1.2.3 Data analysis

Due to the large number of spectra generated by modern mass spectrometers, peptide and protein identification is a complex multi-step process that requires the use of automated search engines[209,210]. While most algorithms compare the experimental m/z ratios of each acquired spectrum to those of a pre-defined database, some are based on *de novo* sequencing[211]. In all cases, each spectrum is matched to a list of possible peptide sequences associated to a defined scoring system, indicating the reliability of the interpretation. In DDA data analysis, the next step consists in matching the peptide against a protein database, manually annotated (UniProtKB/Swiss-Prot[212]) or custom-made[213]. In immunopeptidomics, additional steps are required to ensure that the peptides identified are possible HLA ligands, and efforts have been made to develop user-friendly tools, such as MhcVizPipe, for that purpose[214].

On the other hand, data obtained with DIA strategies impose the use of spectral libraries generated from samples or pools of synthetic peptides analyzed in DDA[215,216]. Spectra from DIA contain a mixture of peptides (as all ions are fragmented and acquired simultaneously)

making their interpretation complex and calling for yet another type of dedicated algorithms and softwares[216]. Fortunately, such computational tools have been developed. Among the most used ones are Skyline[217], Spectronaut[218], and OpenSWATH[219].

2.3.2 Neo-antigens

Neo-antigens are peptides with immunogenic properties, of non-canonical origin, and relevance in a disease setting. Due to their disease specificity, they are actively sought as they might be more easily implemented in diagnostic and therapeutic strategies[220–223]. Many different approaches have been experimented to identify neo-antigens of various nature.

2.3.2.1 Genomic mutations

One of the earliest strategies achieved and still widely used is termed proteogenomics and takes advantage of genomics (DNA sequencing) and transcriptomics (RNA sequencing, ribosome profiling) technologies to build custom databases to help mass spectra interpretation of proteomics and immunopeptidomics datasets. This allows the easy identification of mutated antigens, possibly unique to the disease model or patient sample[205,213]. This approach is particularly relevant in the cancer setting, where mutated antigens are often highly immunogenic.

2.3.2.2 Post-translational modifications

The variety of genes constituting the genome is further diversified by the proteome, as a single gene can lead to several transcripts (alternative promoter, mRNA editing,…). Post-translational modifications (PTMs) are defined as biochemical modifications that occur after the translation of a protein, on one or several of its amino acids, and they add yet another level of complexity to the proteome as they increase its functional diversity[224].

The >1,500 entries in the Unimod database, which aggregates protein modifications for MS, illustrate the manifold nature of PTMs. The presentation of PTMs on peptides bound to HLA molecules can be crucial for T cells responses, but their identification remains deeply challenging[225]. Most of the algorithms used for immunopeptidomics data analysis perform poorly due to the large search space. Besides, one PTM can have the same molecular weight as another or as a combination of others (a phenomenon termed isobaric PTMs) leading algorithms to assign the wrong PTM to a peptide. Only a very resolutive instrument could distinguish between near-isobaric PTMs. The accuracy of measurement is also critical when localizing the residue which carries a PTM[225]. Finally, some modifications are known to arise artifactually, making the characterization of biologically relevant PTMs even more complicated.

Nonetheless, a couple of studies have confidently reported PTMs on HLA-bound peptides in cancer[204,226] but only a portion have been investigated for their T cell recognition[205].

Several post-translationally modified antigens have also been described in T1D but, to date, none of them carries both the evidence of natural processing and presentation and T-cell recognition. The effect of oxidation reactions on the major β-cell autoantigen, GAD has been investigated. They drive the formation of GAD aggregates that are preferably recognized by aAbs of T1D patients, but their possible implication in the formation of oxidized T-cell epitope has not been tested[227,228]. Using *in silico* predictions, synthetic deamidated peptides have been found to preferentially bind to the T1D-associated HLA-DQ8 transdimer (DQA1*05:01-DQB1*03:02) and to be recognized by $CD4^+$ T cells[229]. Later, other deamidated sequences were HLA-eluted from DCs pulsed with pools of autoantigens but their recognition by T cells was not investigated[230]. Citrullination of the ER chaperone BiP (also known as GRP78) has first been identified in islets of NOD mice, at the protein level[231]. Lately, we showed that citrullinated GRP78 peptides are variably recognized by circulating $CD8^+$ T cells (Annex 4)[232]. Lastly, carbonylation has been detected on islet proteins in NOD mice before the onset of the disease[233]. Interestingly, one of these proteins, the prolyl-4-hydroxylase beta subunit (P4Hb), is involved in proinsulin folding and trafficking. Circulating aAbs and $CD4^+$ T cells reactive to P4Hb were observed but as for other studies, the natural generation of modified epitopes has not been investigated.

2.3.2.3 Peptide splicing

The main role of the proteasome is to degrade proteins into peptides but it can also create splice peptides through the ligation of two non-contiguous fragments from the same protein (*cis*-splicing) or from two different proteins (*trans*-splicing) (Figure 17). This process has caught a lot of attention in recent years, and its quantitative contribution to the overall immunopeptidome is still controversial[234–237].

By definition, the sequences of spliced peptides are not templated in the genome, making their identification impossible with classical databases. One of the earliest strategies used to overcome this issue consisted in extending the reference database to include all possible spliced products[234]. Unfortunately, this drastic growth in the size of the reference database compromises usual scoring systems and exponentially increases the chances of finding false positive hits[237]. Later, more complex analysis pipelines were developed to limit the search space and ensure the identification of true spliced peptides[235,236]. The overall existence of *trans*-splice

peptides is still questioned, pleading that the chance for two fragments of two different proteins to enter the proteasome at the same time and to be co-processed is too low[237]. β cells may represent an exception, as the formation of these *trans*-splice peptides may rather take place in insulin granules[238,239]. Altogether, the correct identification of spliced peptides suffers from computational limitations and from the many proteasome-independent alternative sources from which supposedly spliced peptides could come from[237,240]. Several algorithms have emerged, attempting to solve these issues and to allow easy and accurate identification of spliced peptides[241–243] (reference 242 can be found in Annex 5).

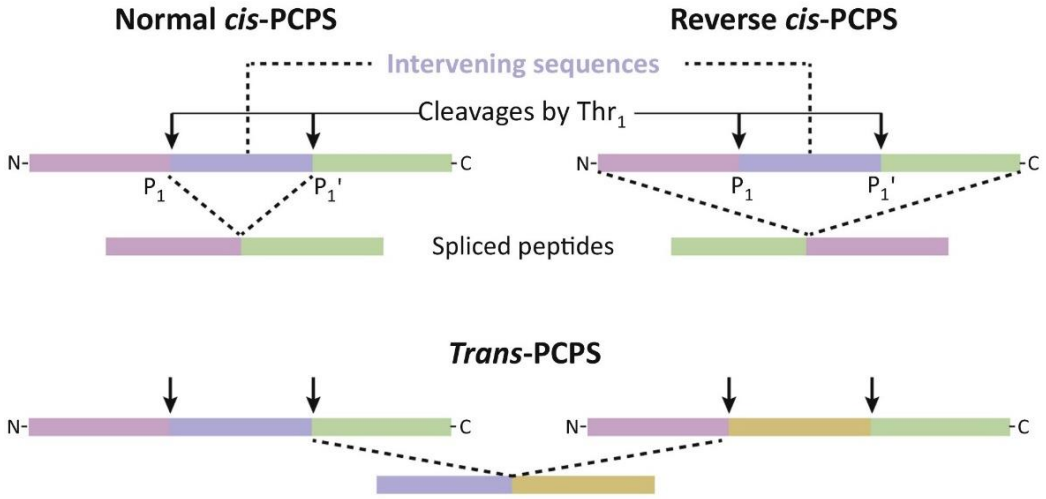


Figure 17. Proteasome-catalyzed peptide splicing (PCPS)
(Top) When the two fragments of the spliced peptide originate from the same substrate, the ligation can occur in the normal order (normal cis-PCPS) or in reverse order (reverse cis-PCPS). (Bottom) When the two fragments originate from two different proteins, the resulting peptide is a trans-PCPS. Thr_1: threonine residue of the proteasome that cleaves the peptide bond. Adapted from[244].

In T1D, only a handful of proteasome-catalyzed *cis*-spliced peptides have been reported and only one was tested for its recognition by circulating $CD8^+$ T cells, the others being allegedly non-binders[188,245]. Interestingly, hybrid insulin peptides (HIPs), which are *cis*-spliced peptides of insulin have first been discovered in NOD mice and since then have also been found in human islets[238,239,246]. These peptides are targeted by both human and murine $CD4^+$ T cells and their formation is likely to occur in secretory granules rather than through the proteasome[247].

2.3.2.4 Other origins

Translation of a transcript outside of the primary open reading frames (ORFs) gives rise to alternative products, that can access the HLA APP pathway and contribute to the peptide repertoire. These cryptic peptides have been reported to stem from long non-coding RNAs, 5'

and 3' untranslated regions, out-of-frame alternative ORFs, intronic retention, codon read-through, and pseudo-genes[240,248–251]. Additionally, ribosome profiling has revealed many unannotated ORFs[252]. A re-analysis of published data considering these possible sources has led to the re-assignment of previously reported spliced peptides to linear peptides from novel ORFs[252], further supporting a central role for these alternative ORFs and also emphasizing the importance of a proper database design[253]. The suitable database should be as small as possible while containing all the sequences that are biologically plausible and relevant in the sample studied.

2.3.3 T1D epitopes

Considering the central role of T cells in T1D pathogenesis, identifying their molecular targets represents a key point in understanding the disease. HLA-ligands that are recognized by T cells are defined as epitopes and considerable efforts have been made to determine the sequence of disease-relevant ones in the context of T1D. In their review, James, Mallone *et al.*[245] have comprehensively assessed the level of evidence for every published peptide in T1D. They have created a grading system based on one hand, on the evidence of natural processing and presentation, and on the other hand, on the source of T cells that recognize the corresponding epitope. They weighted the evidence of natural processing and presentation highly, as the absence of such evidence means that the considered peptide could be a mimotope rather than a natural antigen (Figure 18).

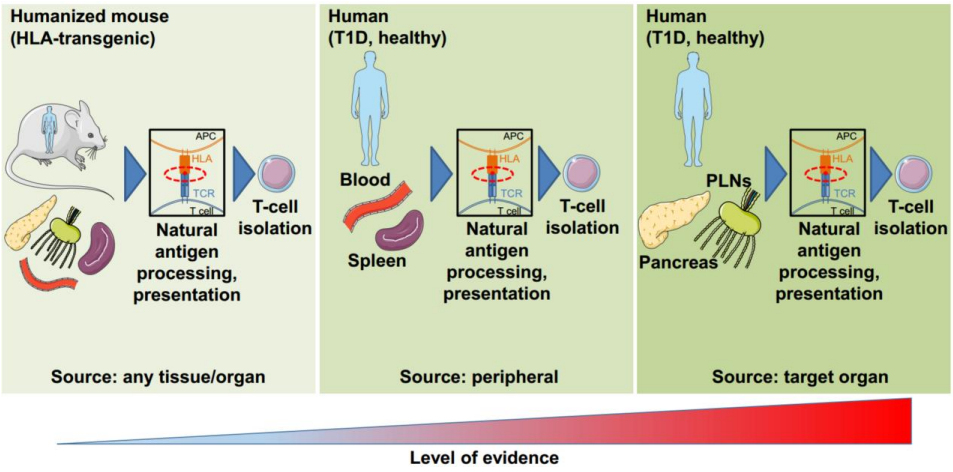


Figure 18. Levels of evidence for islet-derived T-cell epitopes
From [254]

In summary, 16 source proteins for HLA-I epitopes have been described and 5 of them represent most epitopes. Indeed, there are 42 epitopes described for INS, 33 for ZnT8, 18 for GAD, 9 for IA-2, and 8 for glucose-6-phosphatase 2 (IGRP). This review also highlighted a

discovery bias, with the vast majority of epitopes identified in the context of the HLA-A*02:01 allele, leading to limited coverage of other HLAs, despite, for example, HLA-A*24:02 and B*39:06 being predisposing for T1D. The picture is quite similar for HLA-II epitopes with 13 source antigens depicted and 3 of them (INS, GAD, IA-2) being the source of most unique epitopes. These 3 proteins were the earliest antigens found but more recently and based on multiple studies, ZnT8 and IGRP are also emerging as predominant antigens. Among the HLA-II epitopes listed, many are overlapping while they certainly share the same minimal core epitope, this reflects the flexibility allowed by the HLA-II molecules. The discovery of HLA-II epitopes appears to be biased toward HLA-DR4 restriction, again leading to limited coverage of other alleles.

Overall, 41% of HLA-I and 82% of HLA-II epitopes lack evidence of natural processing and presentation, questioning whether they are true disease-associated sequences and emphasizing the need for additional studies. Moreover, only a minority of epitopes (6% for HLA-I, 8% for HLA-II) have been found in T cells infiltrating the islets or from pLNs. Only 4 out of 142 HLA-I epitopes are convincing in all aspects considered ($GAD_{114-123}$, INS_{15-24}, INS_{34-42}, $ZnT8_{186-194}$), and 3 out of 418 HLA-II ($GAD_{115-127}$, $GAD_{274-286}$, INS_{33-47}). Notably, HLA-I- and HLA-II-restricted $GAD_{114-123}$/ $GAD_{115-127}$ and INS_{34-42}/ INS_{33-47} epitopes are overlapping.

3. Benign autoimmunity and loss of β-cell ignorance

3.1 We are all autoimmune

The role of T cells in T1D pathogenesis has been investigated in many studies, both in humans and mice[24]. It is now clear that they are central to the disease development, as exemplified by the observation that NOD mice deprived of T cells are protected from diabetes and that both $CD4^+$ and $CD8^+$ T cells are required to induce the onset[255]. T cells are activated in the pLNs by their cognate antigens, then autoreactive T cells migrate to the pancreas, locally triggering a series of events that results in inflammation (Figure 6).

Due to their low frequency in human blood and the difficulty to access pancreas specimens, the analysis of autoreactive T cells is challenging. Thankfully, technological advances and the development of initiatives dedicated to organ collection, more specifically nPOD[256], have alleviated some of these limitations and permitted substantial progress in the identification of autoimmune T cells and their features.

Using a combinatorial HLA-I multimer (MMr) strategy,which allows the reliable detection of rare events[257–259], our laboratory has recently shown that islet-reactive CD8$^+$ T cells unexpectedly circulate at similar frequencies in the blood of recent-onset T1D patients and in sex- and age-matched healthy individuals[119,188]. The frequency of autoimmune T cells recognizing known antigens (preproinsulin, GAD, IA-2, IGRP, and ZnT8) and novel antigens (UCN3, SCG5, PCSK2, IAPP) is comparable and falls in the range of 1-50/10^6 CD8$^+$ T cells expected for naïve precursors (Figure 19)[119]. Indeed, a large fraction of these islet-reactive CD8$^+$ T cells display a naïve-like phenotype (CD45RA$^+$CCR7$^+$) in both groups of subjects, suggesting that they are not actively involved in the autoimmune processes[188,260]. These observations hold true across HLA-I restrictions[260] and have been confirmed by other groups[261,262].

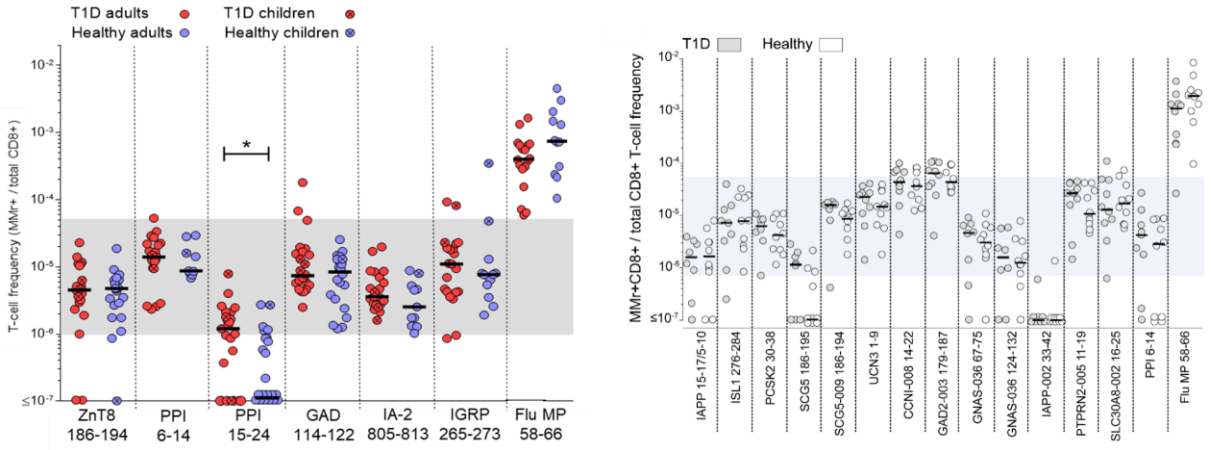


Figure 19. Frequency of islet-reactive CD8$^+$ T cells in T1D patients and healthy individuals

For both for **A.** known antigens and **B.** novel ones, the frequency of recognition by islet-reactive CD8$^+$ T cells is similar in T1D (red) and healthy (blue) blood donors.

In contrast, CD8$^+$ T cells with the same antigenic specificities are enriched in the pancreas of T1D patients as compared to non-diabetic subjects and present a memory phenotype[119,188]. Both CD4$^+$ and CD8$^+$ T cells, that recognize β-cell autoantigens have been reported in islets of at-risk, recent-onset, and longstanding T1D patients[119,188,263–266]. These observations imply that there is a local expression or release of factors that drive T cell migration to the tissue and activation. It remains however unknown whether all the islet-infiltrating T cells are autoreactive, or if bystander-activated T cells are also recruited to the tissue upon inflammation[267,268]. One argument against the presence of non-β-cell antigen-specific T cells within the islets is the requirement for both HLA-I expression and local antigen recognition for CD8$^+$ T cells to home to the pancreas[269].

The similar frequency and phenotype of circulating $CD8^+$ T cells in the blood of T1D and healthy donors prompted us to introduce the concept of a universal state of "benign islet autoimmunity" occurring in all individuals. This hypothesis is reinforced by analogous observations in other autoimmune diseases, such as multiple sclerosis, in which T cells reactive to myelin oligodendrocyte glycoprotein (MOG) and proteolipid protein (PLP) antigens are detected in the same proportion in the blood of patients and healthy individuals[270–272]. This benign autoimmunity stems from an incomplete clonal deletion in the thymus[119,120] and challenges the current dogma on T1D development. Indeed, it was previously suggested that T1D arises from defects in the negative selection of autoreactive T cells, leading to their escape to the periphery and consequent activation and expansion[262,273,274]. The impact of central tolerance mechanisms is likely minimal, as confirmed by the observation that the frequency of circulating T cells reactive to antigens expressed (CHGA, ISL1, SCG5, IA-2, IGRP) or not expressed (GAD, ZnT8) by medullary thymic epithelial cells (mTECs) falls in the same range[119]. In addition, HLA-A2-restricted epitopes are recognized, to the same extent, by $CD8^+$ T cells from HLA-$A2^+$ and HLA-$A2^-$ donors, while only the former is capable of presenting such epitopes[119]. Altogether, the accumulated data imply that the earlier assumption of a disease-related leaky thymic selection may rather be a universal process.

Nevertheless, several limitations should be considered. First, most of these conclusions were drawn from HLA-I MMr-based investigations, which preferentially detect T cells with high TCR affinity and may induce an analysis bias. In a recent study, CD8+ T cells displaying a functional reactivity against self-antigens were uncovered although they are not recognized by MMr[275]. Second, the use of MMr limits the number of markers that can be included to phenotype the MMr-positive T cells. In most cases, only the classical distinction of naïve ($CCR7^+CD45RA^+$), effector memory ($CCR7^-CD45RA^-$), central memory ($CCR7^+CD45RA^-$), and TEMRA (terminally differentiated effector memory T cells re-expression CD45RA; $CCR7^-CD45RA^+$) can be made. The prevalent naïve phenotype of islet-reactive $CD8^+$ T cells may truly encompass less common T-cell subsets that cannot be discriminated when solely using CCR7 and CD45RA. For example, stem memory T cells (Tscm) are increased in T1D patients[276] and 'memory T cells with a naïve phenotype' (Tmnp) arise in response to repeated antigen stimulation, and both of these subtypes share the surface expression of CD45RA+CCR7+ with naïve T cells[277]. Third, these studies remain largely descriptive and do not interrogate the functionality of T cells, which could perhaps be distinguished by a unique status of activation or migration, and/or their sensitivity to regulation. Fourth, additional studies

are required in at-risk donors to elucidate whether the observed lack of difference between $CD8^+$ T cell frequencies between T1D and healthy donors is specific to clinical T1D, while perturbations might exist earlier in stage 1-2 of the disease. Fifth, these studies are focused on $CD8^+$ T cells and comparable ones remain to be carried out on their $CD4^+$ T cells, which could behave differently.

3.2 Accomplice β cells

β cells have always been considered innocent victims of $CD8^+$ T cell attacks in the pancreas. This tenet is currently challenged by accumulating evidence highlighting the active contribution of β cells to their own demise, by increasing their visibility and engaging in a conflicting dialogue with the immune system (Figure 20)[278,279].

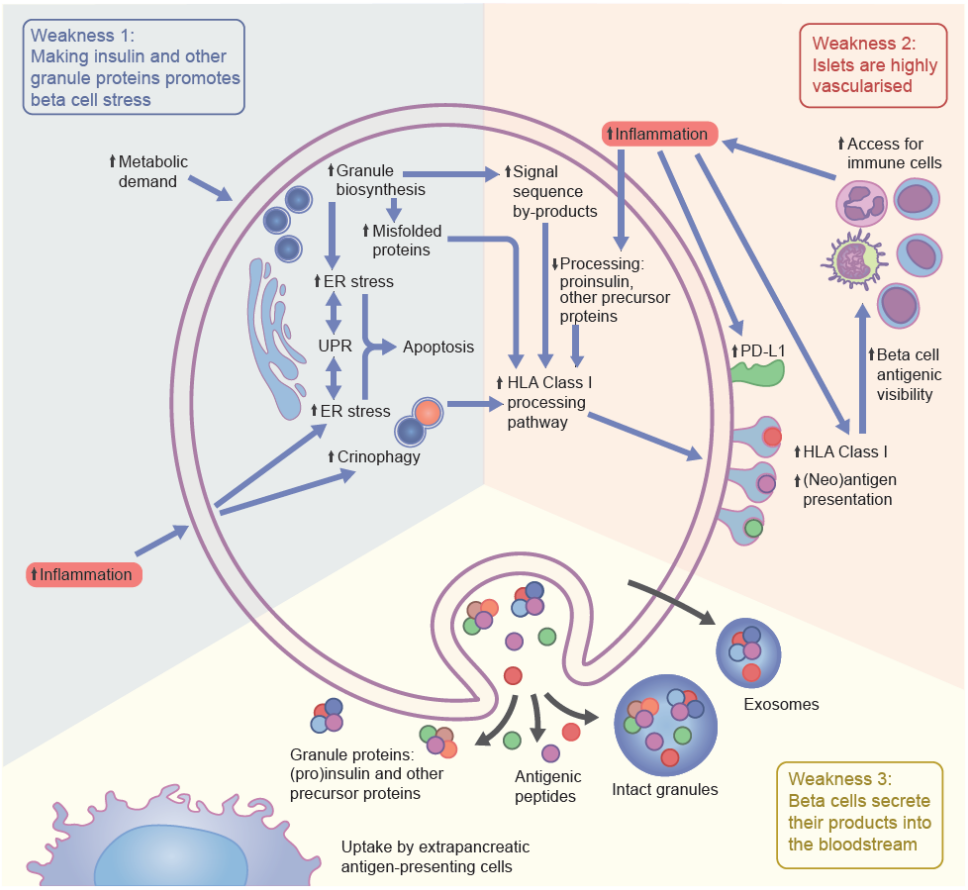


Figure 20. β-cell biology weaknesses promoting autoimmune vulnerability
From [278]

3.2.1 Metabolic β-cell stress

By definition, β cells respond to blood glucose levels by releasing the appropriate amount of insulin. They need to constantly adapt to metabolic changes, and being long-lived virtually non-proliferating cells, they can only achieve this goal by modulating their function,

not their numbers. Upon nutrient uptake, they can increase the synthesis of (pro-)insulin more than 10-fold. This constant adaptation relies on a well-balanced unfolded protein response (UPR), that decreases the protein translation rate, enhances the synthesis of protein-folding chaperones, and increases the degradation of misfolded proteins[280]. If the UPR fails at restoring cellular homeostasis, the sustained ER stress can lead to the local release of cytokines and chemokines that participate in the establishment of inflammation[281] and eventually to the activation of apoptotic pathways[280,282]. Thus, β cells are highly active cells that operate on the edge between physiological and pathological UPR.

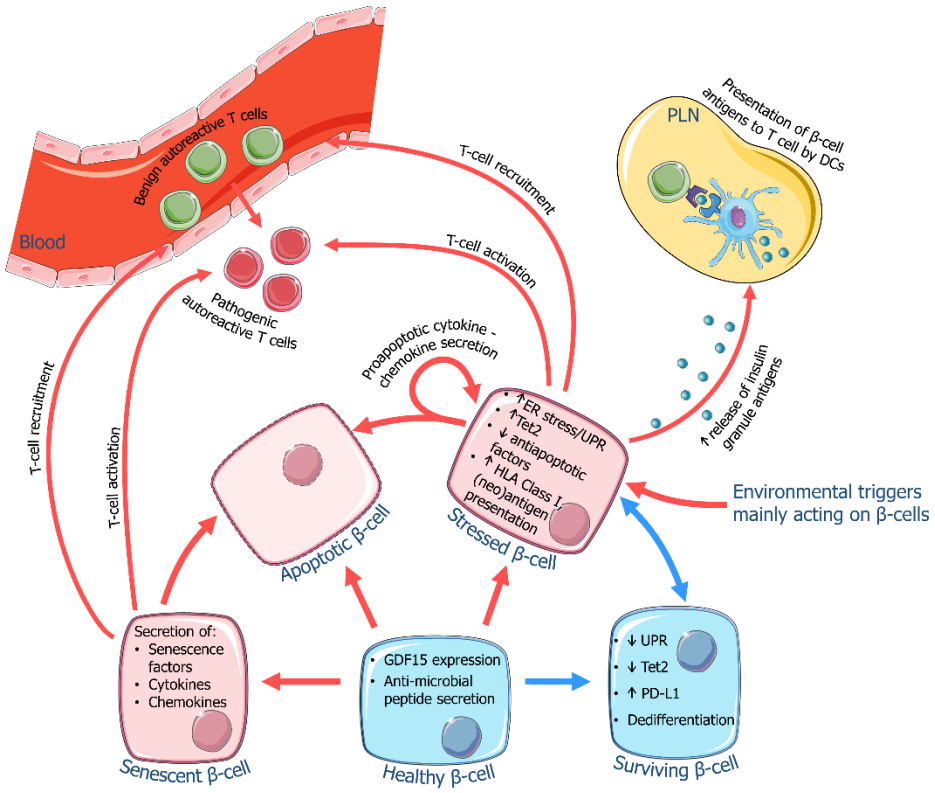


Figure 21. β-cell role in their own demise
Mechanisms by which β cells participate in their homicide (from[279])

3.2.2 β-cell susceptibility to its environment

The delicate metabolic balance that β must maintain makes them particularly sensitive to additional stressors, such as viral infections and cytokines. As mentioned in section 0, CVBs are the main environmental suspect for T1D development. Although most studies investigating the molecular and cellular mechanisms of CVB replication cycle are conducted in cell types other than β cells, the elucidated processes can be extrapolated to β cells to try to understand how such a strong external cue could synergize with ER stress (Figure 21).

The outcomes of CVB infection include intensive viral protein production, impaired cellular calcium homeostasis, and ER membrane modifications[283,284], all of which potentiate ER stress and imbalance the UPR[285]. Besides, CVB concomitantly reduces the anterograde vesicular trafficking of the cells it infects and increases the retrograde trafficking[286], which translates into reduced secretory activity and insulin stores[287], likely inflicting additional cellular stress. Moreover, CVB hijacks the autophagy pathway to promote its replication in vesicular hubs[288–290]. Autophagy is a key process that allows β cells to dispose of old insulin granules and renew their stocks[178] and its impairment certainly participates in cellular stress.

The observation that the neighboring α cells are spared from islet autoimmunity further supports the role of ER stress and environmental factors. Indeed, α cells are more resistant to metabolic stress-induced apoptosis[291] and mount more efficient anti-CVB responses[292].

On a molecular level, the double-stranded (ds)RNA replicative intermediate can be recognized by pattern recognition receptors such as toll-like receptor 3 (TLR3), a retinoic-acid-inducible gene I (RIG-I), and melanoma differentiation-associated protein 5 (MDA5), which potentiate the production of pro-inflammatory cytokines, notably type I IFNs. MDA5 expression is upregulated in human islets upon CVB infection and in recent-onset T1D donors[293]. Moreover, infection of human islets with CVB induces the production of IFN-α[294,295], suggesting that the early blood IFN-α signature of T1D may reflect an anti-viral response[86,296]. While this IFN response limits CVB replication and spreading, it also enhances beta-cell apoptosis (21). Through the IFN-α pathway, dsRNA also dedifferentiates β cells and impairs their insulin secretion[297].

3.2.3 β-cell antigenic visibility

The early IFN-α signature that characterizes T1D could result from CVB infection but also from pDCs and/or neurophils[69,70]. ER stress is further amplified by this inflammatory milieu and, in human islets, β-cell function is hindered[298,299]. In mice, the BDC2.5 clones respond higher to islets in which ER stress has been chemically-induced, which also led to an increase of tissue transglutaminase 2, an enzyme catalyzing deamidations[300]. This observation hints that the heightened immunogenicity could depend on the formation and presentation of neo-antigens on surface HLA-I[300]. One of the histopathological hallmarks of T1D is the hyperexpression of surface HLA-I molecules in the islets and on β cells, meaning that more peptides can be presented at the cell surface[142,188,301,302]. This boosts β-cell visibility and disrupts

the immune ignorance of these cells, making them more vulnerable to the immune system (Figure 20).

Disease development starts with cellular and humoral responses against one autoantigen, and then spreads to other β-cell antigenic targets. The first markers of loss of self-tolerance in T1D are aAbs directed against insulin or GAD, associated with the detection of cognate autoreactive T cells infiltrating the pancreas before clinical diagnosis[266,303]. Autoimmunity is further amplified by the process of "antigen/epitope spreading", in which the immune response specific to a single epitope propagates to other peptides from the same protein or from other proteins with identical cellular origin[304]. This mechanism was first depicted in the experimental autoimmune encephalomyelitis mouse model[305], which is widely used to study multiple sclerosis and was later investigated in T1D[306]. NOD mice were genetically-modified so that their APCs overexpress either IGRP (NOD-IGRP) or proinsulin 2 (NOD-PI). NOD-IGRP mice, which are tolerant to IGRP, develop diabetes and insulitis as the conventional NOD strain. On the contrary, diabetes was prevented in NOD-PI mice, which are tolerant to proinsulin 2, thus providing direct evidence that PI plays the primary role in autoimmune diabetes while immune responses to IGRP lie downstream of it. The situation is less clear in the human disease, even though insulin antigenicity seems central, at least in T1DE1.

Our laboratory has recently identified disease-relevant epitopes using HLA-peptidomics and transcriptomics approaches on an *in vitro* β-cell line exposed to IFN-γ alone or in combination with TNF-α and IL-1β to mimic the inflammatory stress characterizing insulitis[188]. We have provided a catalog of HLA-I-bound peptides naturally processed and presented by β cells, containing known and novel antigens. Our study emphasized the predominance of secretory granule products, among which UCN3, SCG5, and PCSK2, which are important enzymes involved in the insulin maturation pathway. PCSK2 is involved in the first maturation step of many neuroendocrine hormones[307,308] and SCG5 is a chaperone of PCSK2 and is continually degraded along the secretory pathway to prevent the premature activation of PCSK2 by autocatalytic cleavage[309,310]. UCN3 is implicated in the regulation of insulin secretion and the maintenance of β-cell function and mass[311]. Focusing on the most common alleles in the Caucasian population, HLA-A2, and HLA-A3, we have observed that several peptides were repeatedly recognized by circulating $CD8^+$ T cells of T1D and healthy donors[188,260]. In the islets of prediabetic NOD mice, a significant proportion of $CD8^+$ T cells responsive to the orthologous sequences of some of these newly-identified antigens have been found[260]. Furthermore, $CD8^+$ T cells, which recognize a pool of 5 peptides from the above 3 proteins are capable of

transferring diabetes into immunodeficient recipient mice, emphasizing the pathological role of these antigens and their cognate T cells[260]. These data imply that insulin granules might constitute an immunogenic-rich compartment. that contributes to the dialogue between β and T cells and could thus participate in the establishment of autoimmunity[24,312].

3.2.4 β-cell dialogue with the immune system

Like other endocrine cells, β cells secrete their granule content directly into the bloodstream, making β-cell antigens easily accessible to APCs in the draining lymph nodes. This physiological feature could prime islet-reactive T cells at distance from the pancreas[313]. Thus, insulin granules composition might contain key prognostic and diagnostic markers of β-cell stress and would deserve dedicated investigations to understand better their role in T1D onset. Additionally to the content of insulin granules, β cells also secrete intact insulin granules and extracellular vesicles, which could also be immunogenic, and perhaps their immunogenic landscape differs from the one of β cells[314].

The inflammatory microenvironment favors the activation of innate immune cells and increases the vascular permeability, which together with the rich vascularization of pancreatic islets facilitates the migration of macrophages and DCs to the islets. Once there, they can uptake apoptotic β cells or β-cell antigens and then home to the pLNs and prime cognate T cells[315]. In turn, activated T cells attracted by β-cell-released chemokines, migrate to the pancreas where they encounter their targets and kill them[315]. The cytokines secreted by these immune cells, TNF-α and IFN-γ, sustain the inflammatory environment, the high HLA-I expression, and the associated increased β-cell visibility and death. However, this cascade of events requires a clear β-cell fragility to be initiated. In humanized immunodeficient mice, human diabetogenic T cells efficiently transfer the disease only if basal β-cell damage is induced by low-dose streptozotocin[316].

Inflammation also triggers the upregulation of programmed cell-death ligand 1 (PD-L1). This observation was specific to *ex vivo* islet sections from T1D patients and was not made in at-risk, T2D, or non-diabetic individuals[317,318]. PD-L1 expression can also be induced *in vitro*, in a JAK/IRF1-dependent mechanism, on EndoC-βH1 cells, human and mouse pancreatic islets, upon IFN-α or IFN-γ treatment[317]. PD-L1 hyperexpression might reflect an ineffective attempt of β cells to escape T cell lysis.

Altogether, β-cell vulnerabilities might make them more visible to the immune system and play a central role in the onset of autoimmune mechanisms leading to T1D (Figure 22).

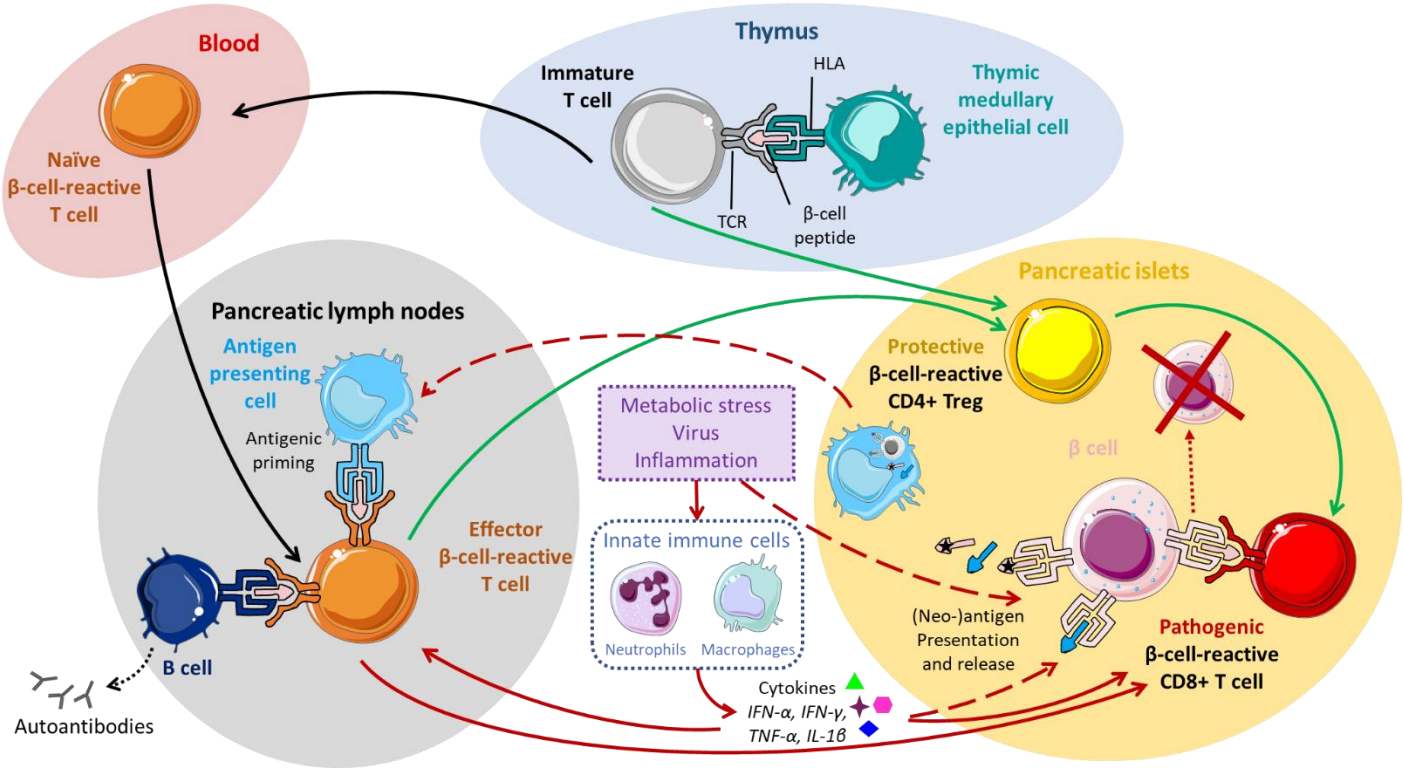


Figure 22. Revised T1D immunopathogenesis

Metabolic stress, viral infection and the resulting inflammation lead to HLA-upregulation at the surface of β cells along with the formation of neo-antigens and the release of cytokines by innate immune cells. The antigens released can be uptaken by APCs on-site or at distance and used to prime T cells in the pLNs.

3.2.5 β-cell targeting therapies

Several therapeutic strategies aimed at restoring insulin secretion by preserving β-cell mass are currently under development or clinical trial. First, the gut hormone glucagon-like peptide 1 (GLP-1) reduces the levels of glucagon while increasing insulin secretion. In T1D patients, this mechanism could stimulate β cells that have escaped the immune attack[319]. In clinical studies, GLP-1 infusion in longstanding T1D patients has led to a 50% decrease in the isoglycemic meal-related insulin requirement, due to increased glucose utilization and a lower glucagon release[320]. Despite the evidence in mice that GLP-1 has anti-apoptotic properties[321,322], the effects observed are likely independent of β cells[320,323]. Nevertheless, GLP-1 remains a promising drug for long-term use in T1D pathology. In another clinical trial, the combination of a GLP-1 analogous (liraglutide) and anti-IL-21 antibody (NCT02443155) led to a significant decrease in C-peptide concentration at 54 weeks, as compared to the control groups (each agent alone and placebo), suggesting that the administration of both drugs ensures a synergistic effect of β-cell preservation and immune modulation[324].

Second, the drug AZD1480, which is a specific inhibitor of JAK1/JAK2, prevents the HLA-I upregulation in murine and human β cells and thus limits the direct interaction with $CD8^+$ T cells and blocks the harmful effect of inflammatory cytokines[325]. In NOD mice, this process allows diabetes reversion and prevention. These investigations prompted clinical trials in recent-onset T1D with other JAK inhibitors. Following Imatinib treatment, C-peptide levels were significantly lowered but this effect was not sustained (NCT01781975)[326]. A possibly more promising drug is Baricitinib, which blocks key signaling pathways in β cells and prevents T cell activation (NCT04774224).

Third, Verapamil, a commonly-used and approved anti-hypertensive and anti-arrhythmic treatment, has shown promising results. Indeed, in a small-scale randomized and double-blinded clinical trial, the oral administration of verapamil for 12 months in addition to the standard insulin therapy in recent-onset T1D has stimulated endogenous β-cell function, lowered needs for exogenous insulin, and reduced hypoglycemic episodes[327]. The INNODIA consortium is currently driving a phase II clinical trial "Ver-A-T1D" (NCT04545151).

Aims of the PhD project

Recent studies in the laboratory have highlighted the existence of a universal state of benign autoimmunity. Taken together with the weaknesses that β cell carry, we have hypothesized that the switch from this benign to pathological autoimmunity rely on an enhanced visibility of β cells to the immune system. My PhD project aims at characterizing the antigens presented by β cells in the events that precede clinical T1D.

1. Currently, CVB is the main environmental factor suspected to trigger autoimmune processes. To gain evidence on the possible causal relationship between CVB and T1D, I took advantage of immunopeptidomics approaches to elucidate the viral HLA-bound peptides presented by a β-cell line that has been infected *in vitro* with two different CVB serotypes. The $CD8^+$ T cell responses have been studied in parallel projects using PBMCs from healthy and T1D adults and children (article #1).
2. CVB infection can lead to the secretion of cytokines, such as IFN-α, which is the earliest cytokine signature observed in the blood of at-risk individuals. The second aim of my work was to detail the immunopeptidome of β cells exposed to IFN-α. Using transcriptomics and computational tools, a dedicated effort was made to identify neo-epitopes, carrying PTMs, deriving from mRNA splice variants or resulting from peptide *cis*-splicing. The role of the proteasome and the immunoproteasome in the formation of neo-epitopes and the generation of a representative set of antigens was investigated by defining the changes in the immunopeptidome of β cells in which these proteasomes were inhibited. (article #2).

Results

Article #1

Coxsackievirus infection of β cells triggers cell lysis but poor CD8⁺ T-cell responses

Alexia Carré[1*], Federica Vecchio[1*], Daniil Korenkov[1], Zhicheng Zhou[1], Paola Apaolaza[2,3], Soile Tuomela[4], Isaac Snowhite[5], Orlando Burgos-Morales[1], Javier Perez-Hernandez[1], Barbara Brandao[1], Georgia Afonso[1], Clémentine Halliez[1], John Kaddis[6], Sally C. Kent[7], Maki Nakayama[8], Joelle Vinh[9], Yann Verdier[9], Raphael Scharfmann[1], Michele Solimena[10], Zuzana Marinicova[10], Janine Sanchez[5], Carmen Bustamante[5], Patricia Gomez[5], Soren Buus[11], the nPOD-Virus Working Group, Sylvaine You[1], Alberto Pugliese[5], Heikki Hyoty[12], Teresa Rodriguez-Calvo[2,3], Malin Flodstrom-Tullberg[4], Roberto Mallone[1,13].

[1]Université de Paris, Institut Cochin, CNRS, INSERM, Paris, France.
[2]Institute of Diabetes Research, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich-Neuherberg, Germany.
[3]German Center for Diabetes Research (DZD), Neuherberg, Germany.
[4]Karolinska Institutet, Department of Medicine HS, Center for Infectious Medicine, Stockholm, Sweden.
[5]Diabetes Research Institute, Leonard Miller School of Medicine, University of Miami, FL, USA.
[6]Department of Diabetes and Cancer Discovery Science, Arthur Riggs Diabetes and Metabolism Research Institute, Beckman Research Institute, City of Hope, Duarte, CA, USA.
[7]University of Massachusetts Medical School, Diabetes Center of Excellence, Division of Diabetes, Department of Medicine, Worcester, MA, USA.
[8]Barbara Davis Center for Diabetes, University of Colorado School of Medicine, Aurora, CO, USA.
[9]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS UMR8249, Paris, France.
[10]Paul Langerhans Institute, Technical University Dresden, Germany.
[11]Panum Institute, Department of International Health, Immunology and Microbiology, Copenhagen, Denmark.
[12]University of Tampere, Faculty of Medicine and Life Sciences, Tampere, Finland.
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, Paris, France.

*A.C. and F.V. contributed equally to this work.

Corresponding author: Roberto Mallone, MD PhD - INSERM U1016 Cochin Institute, G.H. Cochin-Port-Royal, Bâtiment Cassini - 123, boulevard de Port-Royal – F-75014 Paris, France. Phone: +33-1-76.53.55.83 E-mail: roberto.mallone@inserm.fr.

Abstract

Coxsackievirus B (CVB) infection of pancreatic β cells is associated with β-cell autoimmunity. However, anti-CVB T-cell responses are poorly characterized. We identified the viral peptides presented by CVB-infected human β cells and the corresponding $CD8^+$ T cells in CVB-seropositive individuals. CVB efficiently lysed β cells and downregulated HLA Class I expression, which led to the presentation of only few, selected HLA-bound viral peptides. Circulating $CD8^+$ T cells recognized only a fraction of these peptides, and only another sub-fraction was targeted by effector/memory T cells that expressed the exhaustion marker PD-1. Infected β cells were more efficiently lysed by CVB than by anti-viral T cells. A CVB epitope cross-reactive with the GAD β-cell antigen was identified. We conclude that β-cell destruction is primarily induced by CVB itself, possibly in the context of poor immune protection, lending rationale to CVB vaccination trials for type 1 diabetes prevention. CVB-reactive $CD8^+$ T cells provide biomarkers to follow response to infection and vaccination along the natural history of disease.

INTRODUCTION

Environmental factors weigh heavier than genetic predisposition in the pathogenesis of type 1 diabetes (T1D), but remain elusive. Moreover, the increasing prevalence of neutral and protective HLA Class II haplotypes in the T1D population over the last decades (1) and the mounting incidence of T1D (2) suggest that environmental triggers are gaining importance. Such triggers are likely to exert their role very early in life, as most children who subsequently develop T1D experience their first seroconversion for islet auto-antibodies (aAbs) before the age of 2 (3, 4).

Infections by Enteroviruses such as Coxsackieviruses B (CVBs) and Echoviruses have long been suspected as possible triggers for T1D (5). Key features relevant to this hypothesis are their high prevalence (>95% of the population being seropositive) and high incidence in infants and toddlers, their oro-fecal transmission, and their capacity to infect β cells, which express the Coxsackievirus and Adenovirus receptor (CAR) and decay-accelerating factor (DAF) (6) required for CVB entry. CVBs are held as plausible candidates based on association studies in prospective cohorts of genetically at-risk children. These studies documented a temporal correlation between serological evidence of infection by some CVB serotypes (e.g. CVB1) and aAb seroconversion, while infection with other serotypes (e.g. CVB3, CVB6) were inversely correlated in some studies (7), possibly reflecting cross-reactive immunity against infection by T1D-associated serotypes. Another recent study (8) relying on fecal metagenome sequencing documented that the predisposing effect of CVB infections (mostly CVB4 in this study) is on islet autoimmunity (i.e. aAb seroconversion) rather than on its subsequent progression to clinical T1D; and that such predisposing effect is linked to prolonged infections with the same serotype rather than to short, independent infections. These prolonged infections, witnessed by the protracted shedding of the same CVB serotype in sequential stool samples, are indicative of viral persistence, which may underlie more virulent strains or a less effective anti-viral immunity.

In parallel, spatial associations have been highlighted in histopathological studies on pancreas specimens (9-11). Positive staining for viral protein (VP)1 was more prevalent in T1D than healthy donors and co-localized with immune infiltration and HLA Class I (HLA-I) hyper-expression, which are the two hallmarks of T1D. Moreover, only a minority of islets with residual β cells are $VP1^+$, suggesting low-grade persistent infections years after seroconversion. Whether these associations underlie a cause-effect relationship remains unsettled. T1D primary prevention trials based on a pentavalent inactivated vaccine targeting the 5 more prevalent CVB1-5 serotypes (12) are therefore advocated (11). Besides protecting from the established

rare but serious complications of CVB infection (e.g. myocarditis, encephalitis, meningitis), such trials would directly test whether preventing CVB infection affords protection from islet autoimmunity and subsequent T1D. Several gaps in knowledge however exist about the natural immune response against CVBs. Available data is limited to serological studies, with one report suggesting that children who subsequently develop early anti-insulin aAbs lack anti-VP1 neutralizing antibodies (although a simple lack of CVB exposure was not ruled out in this study) (13). In line with the recent data from the TEDDY study (8), this could predispose to infections at higher viral loads, which may favor viremia and dissemination to the pancreas. Whether the lack of passive protection by maternal anti-CVB antibodies may also play a role remains controversial (7, 13).

$CD8^+$ T cells play a key role in the clearance of viral infections through cytotoxic lysis of infected cells. Indeed, infection rapidly leads to intracellular processing of viral proteins and surface presentation of HLA-I-bound peptides for T-cell recognition (14). Against this background, it is unknown whether islet destruction and triggering of autoimmunity reflect a direct β-cytolytic effect of viral infection or an indirect effect of anti-viral T cells against infected β cells. Furthermore, the lack of knowledge about the CVB peptides naturally processed and presented by infected β cells and recognized by $CD8^+$ T cells hampers our possibility to understand this process and to follow response to infection and, eventually, vaccination. We therefore set forth to fill these gaps by identifying the HLA-I-bound viral peptides presented by infected β cells and by using these peptides to track CVB-reactive $CD8^+$ T-cell responses.

RESULTS

CVB-infected β cells downregulate surface HLA-I expression and present only few selected HLA-I-bound viral peptides.

To identify the CVB peptides naturally processed and presented by β cells, we infected the ECN90 β-cell line (15), which expresses the highly prevalent HLA-A*02:01 and -A*03:01 alleles (HLA-A2 and -A3 from hereon) and the CVB receptors CAR and DAF (data not shown). *In-vitro* infection with either CVB3 or CVB1 was highly efficient, leading to expression of VP1 and double-stranded (ds)RNA in >70% of cells within 4 h (Fig. 1A, Fig. S1A), and to increased β-cell lysis at 8 h. For immunopeptidomics experiments, we selected the 6-h time point, which corresponds to the infection plateau preserving viability. Notably, infection down-regulated HLA-I expression (Fig. 1B-C).

Three preparations of ECN90 β cells (96-250×10^6 cells/each) infected with either CVB3 or CVB1 (Fig. S1B) were subsequently analyzed after purification of peptide (p)HLA-I complexes, peptide elution, liquid chromatography – tandem mass spectrometry (LC-MS/MS), and bioinformatics assignment of MS/MS spectra (15), using the amino acid (aa) sequences translated from RNA-sequencing of the CVB1 and CVB3 strains used. A total of 17 unique CVB peptides were identified, which, accounted for a minor fraction (3.4%; range 2.1-3.7%) of the total 8-12 aa peptide display (Fig. 1D-E). The β-cell peptides identified in parallel did not yield any novel hit compared to our previous report (15) (not shown). A search for putative transpeptidation products generated by the fusion of CVB and β-cell peptide fragments using our previous scripts (15) did not retrieve any hit.

Only 3/17 peptides (Fig. S2) mapped to the larger structural capsid protein P1 (VP1-VP4); 6/17 located in the non-structural protein P3 (3A-D). The smaller non-structural P2 (2A-C) protein comprised most (9/17) of the eluted peptides, with 3/9 peptides eluted from both CVB3- and CVB1-infected cells. The percent number of eluted peptides mapping to P2 were significantly more than expected (53% vs. 26%), opposite to the few peptides mapping to P1 (14% vs. expected 35-39%; Fig. S3A). Overall, the eluted peptides were largely conserved across the 6 serotypes, with 5/17 (29%) overlapping between CVB1- and CVB3-infected β cells.

This list of 18 peptides was complemented: *a)* by searching for length variants mapping to the same region of the eluted peptides (15) and displaying better HLA-A2/A3-binding score (n=6; indicated in italics in Fig. 1E); *b)* by an *in-silico* prediction of HLA-A2/A3-restricted peptides from representative strains of all CVB serotypes (n=28; Fig. 1F). Of note, $CVB1_{1356-1364}$/$CVB3_{1359-1367}$ (not listed in Fig. 1F) was the only HLA-I-eluted peptide that was also predicted *in silico*, supporting the complementarity of the two strategies. Although the degree

of conservation across serotypes remained high, this *in-silico* search also added some serotype-specific peptides, e.g. $CVB5_{169\text{-}177}$, $CVB5_{760\text{-}768}$, $CVB5_{945\text{-}953}$.

Minor sequence homologies with β-cell proteins (i.e. IA-2, IGRP, ZnT8) were noted for some CVB peptides (Fig. 1E-F). Furthermore, the eluted $CVB1_{1132\text{-}1140}$ (KILPEVKEK) and $CVB3_{1135\text{-}1143}$ (KILPEVREK) displayed a high homology with a $GAD_{272\text{-}280}$ peptide (KMFPEVKEK; conserved aa underlined) eluted from previous β-cell preparations (15). Besides displaying a higher homology for $GAD_{272\text{-}280}$, $CVB1_{1132\text{-}1140}$ proved to be a better epitope in preliminary T-cell analyses (not shown) and was therefore retained for further studies.

We subsequently focused on the 14 HLA-eluted peptides predicted to be restricted by the most prevalent HLA-A2 and -A3 allotypes. HLA-A2/A3 binding was validated *in vitro* (Fig. S3B-D), leading to the final selection of 38 candidates for T-cell studies (Fig. 1E-F): 10 from immunopeptidomics experiments and 28 from *in-silico* predictions; 25 HLA-A2-restricted and 13 HLA-A3-restricted.

Collectively, these results show that CVB-infected β cells downregulate surface HLA-I expression and present only few selected HLA-I-bound viral peptides, which significantly overlap between CVB1- and CVB3-infected cells, are largely conserved across serotypes and mostly map to the P2 protein.

Limited recognition of CVB peptides by circulating $CD8^+$ T cells of seropositive individuals.

We next tested whether the candidate epitopes identified were recognized by circulating $CD8^+$ T cells, using frozen/thawed peripheral blood mononuclear cells (PBMCs) from CVB-seropositive healthy donors (Table S1) and combinatorial HLA-A2 or HLA-A3 multimer (MMr) assays (15-18). The reproducibility was satisfactory for measuring both MMr^+ T-cell frequencies and naïve vs. effector/memory phenotypes (coefficient of variation 21-27%; Fig. S4). As these assays can probe up to 15 MMr specificities per sample, we screened candidate epitopes in 3 panels: two panels for HLA-A2-restricted candidates (n=12+12; Fig. 2A-B) and one panel for HLA-A3-restricted ones (n=12; Fig. 2C). T-cell frequencies visually segregated into two groups, above or below 1 $MMr^+/10^5$ total $CD8^+$ T cells. This cut-off was therefore selected to retain candidate epitopes, as it corresponds to frequencies above the $1\text{-}10/10^6$ typically observed for most peptide-reactive naïve $CD8^+$ T cells (15-17, 19). Two additional peptides (HLA-A2-restricted $CVB1\text{-}2\text{-}4\text{-}5\text{-}6_{271\text{-}279}$ and HLA-A3-restricted $CVB1_{1356\text{-}}$

$_{1364}$/CVB3$_{1359-1367}$) were retained because, although only a subgroup of donors harbored cognate T-cell frequencies above this cut-off, the MMr$^+$ cells of these donors displayed a predominant (>60%) effector/memory phenotype (CD45RA$^+$CCR7$^-$, CD45RA$^-$CCR7$^+$ or CD45RA$^-$CCR7$^-$). Conversely, only some (6/11) peptides returning high cognate T-cell frequencies displayed a similarly predominant effector/memory phenotype. Overall, 7 HLA-A2-restricted candidates (0/4 HLA-eluted and 7/20 *in-silico* predicted) and 6 HLA-A3-restricted candidates (4/6 HLA-eluted and 2/6 *in-silico* predicted) were retained. Given the high homology between the HLA-A3-restricted CVB1$_{1356-1364}$/CVB3$_{1359-1367}$ (KINMPMSVK) and CVB2$_{1361-1369}$ (KVNMPMSVK), only the former (HLA-eluted) peptide was further considered.

These 12 peptides underwent a second T-cell validation round using more stringent gating and selection criteria. Representative dot plots obtained for these HLA-A2 and HLA-3 MMr panels are shown in Fig. S5-S6. Results are summarized in Fig. 3 and detailed in Table S2. The frequency cut-off was here set at 5 MMr$^+$/10^6 total CD8$^+$ T cells to account for the more stringent gating criteria applied and for the exclusion of samples with <3 MMr$^+$ cells counted to avoid any undersampling bias. Most peptides (5/7 for HLA-A2, 4/5 for HLA-A3) were finally validated as relevant targets based on their high cognate T-cell frequency. Of further note, HLA-A3$^+$ donors were recruited at two different sites (Paris and Miami) to account for any bias related to environmental exposure, with similar T-cell frequencies detected in the two subgroups (Fig. 3C).

The HLA-A3-restricted GAD$_{272-280}$ peptide homolog to CVB1$_{1132-1140}$ was further included at this stage to assess any potential cross-reactivity. To this end, we employed our previous assay modification (18) in which the MMr pair loaded with CVB1$_{1132-1140}$ (PE/CF594- and BV711-labeled) shared the BV711 fluorochrome with the MMr pair loaded with GAD$_{272-280}$ (PE- and BV711-labeled), thus allowing selectively gating of CD8$^+$ T cells recognizing only the CVB1 or GAD peptide, or both. Most T cells displayed cross-recognition of the CVB1$_{1132-1140}$/GAD$_{272-280}$ homologous peptides, as shown by co-labeling with the two MMr pairs (Fig. S6C) and by the higher frequencies of CVB1$_{1132-1140}$/GAD$_{272-280}$ MMr$^+$ cells vs. those recognizing either peptide alone (Fig. 3C).

Overall, 3/10 (30%) HLA-eluted peptides were finally validated as robust T-cell targets. This recognition of only a subfraction of naturally processed and presented peptides did not reflect an MS selection bias, as the same was observed with *in-silico* predicted peptides (7/26, 27% validated as T-cell targets). Moreover, as observed in T-cell screening experiments, even when focusing on these more selected epitopes that are both immunodominant (i.e. high cognate CD8$^+$ T-cell frequencies) and immunoprevalent (i.e. recognized in the majority of donors), the

effector/memory fraction was <50% in most individuals, including for highly immunodominant epitopes (e.g. HLA-A3-restricted $CVB5_{760-768}$). This contrasted with the higher frequency and near-complete effector/memory phenotype of control Influenzavirus-reactive $CD8^+$ T cells. Exceptions were noted for HLA-A3-restricted immunodominant $CVB1_{1356-1364}$/$CVB3_{1359-1367}$-reactive T cells and, to a lesser extent, $CVB1_{1132-1140}$/$GAD_{272-280}$-cross-reactive T cells (Fig. 3C-D). Overall, CVB-reactive $CD8^+$ T-cell frequencies roughly correlated with the percent effector/memory fractions observed (Fig. 3E-F), suggesting that higher frequencies reflected prior *in-vivo* expansion.

Collectively, these data show that only a fraction of the few CVB peptides displayed by infected β cells is recognized by circulating $CD8^+$ T cells of seropositive individuals, and that only another sub-fraction is targeted by predominantly effector/memory T cells. Moreover, $CVB1_{1132-1140}$-reactive T cells cross-recognize a homologous $GAD_{272-280}$ sequence.

CVB-reactive $CD8^+$ T cells display similar features in the blood of adults and children.

One reason for the limited circulating CVB-reactive $CD8^+$ T-cell responses observed in seropositive healthy adults could be the long delay since CVB exposure, as these infections are more frequent during childhood. We therefore analyzed these responses in HLA-$A2^+$ CVB-seropositive children, with or without T1D, recruited at two different sites (Miami and Paris; Table S3), using the HLA-A2 MMr panel selected in the validation round. Given that PBMC numbers are more limited and vary according to age for pediatric donors, we preliminarily defined the coefficient of variation (CV) of the T-cell frequency measurement when decreasing PBMC numbers are sampled (Fig. S7A-B). The minimal total $CD8^+$ T-cell count maintaining the CV below 35% was thus defined for each epitope, and used to exclude pediatric samples with insufficient counts, thus limiting undersampling bias. As observed for CVB-seropositive healthy adults, children recruited at the two sites did not display major $CD8^+$ T-cell differences and were therefore analyzed altogether (Fig. S7C-D). No difference was observed between T1D and healthy children in terms of T-cell frequency, barring a marginally higher frequency of CVB1-2-4-5-$6_{271-279}$-reactive $CD8^+$ T cells in the T1D group. The effector/memory fraction of $CVB5_{945-953}$-reactive $CD8^+$ T cells was higher in healthy children. As in adults, Influenzavirus-reactive $CD8^+$ T cells displayed higher frequencies and effector/memory fractions than CVB-reactive ones. Importantly, the limited magnitude of T-cell responses was similar between healthy children and adults (see Fig. 3), ruling out a waning effect from remote CVB exposure.

CVB-reactive CD8$^+$ T cells are found in pancreatic lymph nodes and display an exhausted phenotype.

A second reason for the limited circulation of CVB-reactive CD8$^+$ T cells could be sequestration in lymphoid organs. We therefore searched for these T cells in the spleen of nPOD organ donors (Table S4-S5). Given the small set of samples available, this analysis did not aim to detect differences across T1D, aAb$^+$ and non-diabetic donor groups, but rather to test whether the frequency and/or effector/memory phenotype was different compared to blood. Despite the limited T-cell sampling imposed by the low cell numbers available, cognate T cells were detected in most donors for HLA-A2-restricted CVB1-2-3-4-5-6$_{1246\text{-}1254}$ and HLA-A3-restricted CVB1$_{1356\text{-}1364}$/CVB3$_{1359\text{-}1367}$ (Fig. 4A). Although frequencies were not higher than in blood, effector/memory fractions accounted for 90-100% of these splenic T cells (Fig. 4B), suggesting lymphoid sequestration of the antigen-experienced pool.

For HLA-A3$^+$ T1D case #6480 displaying high CVB1$_{1356\text{-}1364}$/CVB3$_{1359\text{-}1367}$ reactivity (Table S4), available pancreatic lymph node (PLN) cells expanded *in vitro* retrieved CD8$^+$ T cells reactive to multiple CVB epitopes (Fig. S8A-D). We therefore searched for CVB-reactive T cells in the PLNs of nPOD cases, and in spleen and PBMC samples when available (Table S4-S5; representative staining in Fig. S9). CD8$^+$ T cells reactive to a given CVB epitope displayed similar frequencies across tissues (Fig. 4C). In all cases, an effector/memory phenotype was observed (Fig. 4D). Most of these T cells displayed an exhausted (PD-1$^+$CD25$^-$) phenotype (Fig. 4E), which was instead not observed in control Influenzavirus-reactive CD8$^+$ T cells (Fig. S8E). CVB-reactive CD8$^+$ T cells were instead absent in the pancreas, as assessed by CVB peptide recall assays on carrier T cells transduced with T-cell receptors (TCRs) sequenced from islet-infiltrating CD8$^+$ T cells of T1D donors (20) (data not shown). TCR sequencing of MMr$^+$CD8$^+$ T cells highlighted the presence of expanded CDR3β clonotypes for both CVB1$_{1356\text{-}1364}$/CVB3$_{1359\text{-}1367}$ (Fig. 4F) and, to a lesser extent, CVB1$_{271\text{-}279}$ (Fig. S8F), confirming prior *in-vivo* expansion of these T cells, in agreement with the observed effector/memory phenotype. While some expanded clonotypes were shared across tissues from the same donor, no public clonotypes (i.e. shared across individuals) were detected. Expanded T-cell clonotypes sequenced from PBMCs of seropositive donors were also largely private (Fig. S10).

We then searched for the 2 expanded CDR3β clonotypes detected in T1D nPOD case #6211 in a TCR database (clonesearch.jdrfnpod.org) compiled from polyclonal CD8$^+$ T cells isolated from nPOD LN, spleen and blood samples (21), which includes samples from this particular donor. Intriguingly, these 2 expanded CVB1$_{1356\text{-}1364}$/CVB3$_{1359\text{-}1367}$-reactive CDR3β clonotypes were found with a 100% nucleotide match in CD8$^+$ T cells from PLNs, but neither from spleen

nor from blood. This donor was a 24-year-old woman with 4 years of T1D duration who carried the expected HLA-A3 restriction and whose pancreas had previously been reported as positive for VP1 immunohistochemistry (22), Enterovirus qPCR and proteomics. These clonotypes accounted for a remarkable 4.0% and 59.4% of total $CD8^+$ T cells.

Collectively, these results show that CVB-reactive $CD8^+$ T cells are found in PLNs of T1D donors and harbor expanded clonotypes and an exhausted PD-1^+ phenotype.

Lysis of CVB-infected β cells is mostly viral-mediated.

We next asked whether the limited CVB epitope display observed in infected β cells may translate into cytotoxicity upon contact with CVB-reactive $CD8^+$ T cells. To this end, an immunodominant TCR clonotype reactive to the HLA-A3-restricted $CVB1_{1356-1364}$/$CVB3_{1359-1367}$ epitope (Fig. 5A) was re-expressed into carrier cells. TCR-transduced 5KC T cells exposed to CVB3-infected β cells upregulated expression of a fluorescent NFAT-induced ZsGreen reporter (23), indicating display of the cognate peptide (Fig. 5B). A parallel dose-response standard curve using peptide-pulsed non-infected β cells estimated the concentration of CVB peptide naturally processed and presented by infected β cells to 2.7 nM. Thus, despite being focused on few peptides, the natural CVB antigen presentation is quantitatively efficient. Cytotoxic primary $CD8^+$ T cells were then transduced with the same TCR and exposed to infected β cells (Fig. 5C). While increasing MOI enhanced β-cell lysis in the absence of T cells, the addition of T cells did not increase lysis at the highest 50 MOI used. At lower 10 MOI, T-cell-mediated lysis was evident only at low (1:1) effector:target ratios. Higher (5:1) ratios resulted instead in decreased lysis, suggesting that T-cell killing may exert a protective effect by limiting viral spreading in the culture well.

Collectively, these results indicate that, despite a quantitatively significant viral peptide display, CVB-reactive $CD8^+$ T cells are marginally cytotoxic against infected β cells, which are instead mostly lysed directly by CVB.

DISCUSSION

Despite lack of evidence for a causal relationship, CVB infections are gaining credit as candidate environmental triggers for T1D based on their temporal correlation with aAb seroconversion (8) and on detection of pancreatic hallmarks of infection predominantly in T1D donors. These associations have prompted the development of a pentavalent CVB vaccine covering the most prevalent CVB1-5 serotypes (12), which is currently in a phase I trial in healthy CVB-seronegative adults (NCT04690426). However, several gaps in knowledge need to be addressed to better inform the development of these trials: what are the features of anti-CVB T-cell responses? How to monitor vaccination efficacy through T-cell biomarkers? We focused on $CD8^+$ T cells, which are key players in both anti-viral responses and in the autoimmune attack against β cells. Two different, non-mutually exclusive scenarios could explain how CVB triggers T1D (5). First, weak anti-CVB immune responses could favor viral spreading and persistence, resulting in an enhanced, primary CVB-mediated lysis of infected β cells. Second, an opposite scenario of strong anti-viral responses could lead to secondary lysis of infected β cells by viral-reactive T cells. In both cases, the final outcome would be the release of self-antigens in an inflammatory environment that may initiate/amplify autoimmunity, but the implications for vaccination strategies are different. While CVB vaccination may be beneficial to enhance weak immune responses and limit viral spreading, the second scenario may invite caution, as boosting the anti-viral T-cell response could trigger destruction of infected β cells. Moreover, a third scenario of molecular mimicry between CVB and endogenous β-cell peptides is also possible, as it could turn a physiological anti-viral T-cell response into pathological autoimmunity.

The current study fills some of these gaps with several unexpected observations. First, an efficient immune escape mechanism mounted by CVB is down-regulation of surface HLA-I on infected β cells, as reported for other cell types (24), thus leading to the presentation of very few, selected CVB peptides. A minority of them mapped to the VP1-4 structural proteins that are used in the pentavalent vaccine under trial (12). Altogether, this makes the possibility that CVB vaccination may trigger an anti-viral response against β cells unlikely. Moreover, this vaccine does not comprise the non-structural $CVB1_{1132\text{-}1140}$/$CVB3_{1135\text{-}1143}$ epitope that we found cross-reactive with the endogenous β-cell peptide $GAD_{272\text{-}280}$. This peptide maps to a larger GAD region whose reactivity was previously reported to segregate with T-cell responses to the homologous CVB region in PBMC proliferation assays (25), suggesting the possibility of epitope mimicry and T-cell cross-reactivity. However, a subsequent study (26) using $CD4^+$ T-cell clones dismissed this possibility. We now show that CVB/GAD cross-recognition is at play

for $CD8^+$ T cells, with the large majority of them reactive to both epitopes. To better understand the pathogenic mechanisms of CVB infection, it will be important to rule out that these cross-reactive T cells do not lyse non-infected β cells. This is likely to be the case, as the $GAD_{272-280}$ epitope is poorly presented by β cells (15).

Second, the $CD8^+$ T-cell responses mounted against the CVB peptides identified are even more limited in terms of number of targets and of effector/memory responses, with some targets recognized by exhausted $PD\text{-}1^+$ T cells, as observed in other settings of persistent viral infection (27, 28). These observations provide some rationale to boost immune memory through vaccination. Although CVB-reactive $CD8^+$ T cells were capable of lysing infected β cells, CVB-mediated lysis was predominant even at low MOI, suggesting that the autoimmune trigger predominantly relies on the direct lytic effect of the virus releasing islet antigens rather than on anti-viral cytotoxic T cells. Together with the poor viral peptide display of infected β cells, this suggests that vaccination is unlikely to enhance β-cell destruction.

Third, by combining immunopeptidomics and *in-silico* predictions, we identified several other epitopes that, although not presented by β cells, were capable of eliciting $CD8^+$ T-cell responses. Overall, we identified two epitopes for each of the HLA-A2 and HLA-A3 restrictions considered displaying a particularly high immunodominance and immunoprevalence: HLA-A2-restricted $CVB1_{271-279}$ and $CVB1_{1246-1254}$ and HLA-A3-restricted $CVB1_{1356-1364}$ and $CVB1_{1132-1140}$. One reason for their preeminence may lie in their conservation across CVB serotypes (Fig. S11). Their more limited homology with the poliovirus serotypes included in the inactivate polio vaccine rules out vaccine-induced cross-reactivity. Importantly, a larger panel of epitopes mapping either to structural VP1-4 proteins (e.g. HLA-A2-restricted $CVB1_{169-177}$, $CVB1_{271-279}$; HLA-A3-restricted $CVB5_{760-768}$) or to non-structural P2-P3 proteins (e.g. HLA-A2-restricted $CVB1_{1246-1264}$, $CVB1_{1671-1679}$; HLA-A3-restricted $CVB1_{1132-1140}$, $CVB1_{1356-1364}$, $CVB3_{1765-1774}$) could be used to differentially monitor T-cell responses to natural infection (where both structural and non-structural epitopes should be targeted) vs. those elicited by vaccination (with only structural epitopes recognized), analogous to what is routinely done for hepatitis B virus by testing anti-HBsAg vs. anti-HBcAg serology. This distinction cannot be made with current CVB serology assays.

High frequencies of circulating CVB-reactive $CD8^+$ T cells were noted in some cases, which correlated with higher effector/memory fractions and with expanded private clonotypes. However, the effector/memory phenotype was rather limited in several donors, suggesting a poor memory generation or maintenance over time, which could favor the prolonged infections associated with the development of islet autoimmunity (8), as suggested by serological studies

(7, 13). This poor memory does not seem to reflect a long delay since CVB exposure, as T-cell frequencies and effector/memory fractions were similar between children and adults. It will be important to directly address this possibility by studying longitudinal PBMC samples taken before and after CVB encounter. The other relevant observation is that CVB-reactive populations displayed a near-complete effector/memory phenotype in the spleen and PLNs, suggesting that the long-term effector/memory fraction may rather reside in secondary lymphoid organs. Nonetheless, it may be possible to detect this fraction in transit through the blood when sampling time points closer to CVB exposure.

The similar circulating CVB-reactive T cells observed in T1D and healthy donors was not unexpected, as it may reflect the late stage in disease history sampled with these studies. However, the high dominance (4-59%) of 2 specific clonotypes within the PLN $CD8^+$ T cells of T1D nPOD case 6211 suggest that these responses are more efficiently elicited at the local level in the disease setting, possibly in the context of higher CVB loads gaining access to the pancreas. This possibility is in line with the "polio hypothesis", which proposes that infections at high viral loads favored by weak anti-viral immunity may similarly underlie the association between poliovirus (spreading to motor neurons) and its severe complication poliomyelitis, and between CVB (spreading to the pancreas) and T1D. This hypothesis may also explain the apparent epidemiological paradox of low CVB circulation in areas and time periods of high T1D incidence (29, 30), reflecting poor antiviral immunity of the background population (31). CVB vaccination may obviate this problem, as it has been the case for polio vaccines.

The surface HLA-I downregulation induced in CVB-infected β cells decreased the presentation not only of CVB peptides, but also of endogenous ones. Indeed, only 152-206 8-12mer peptides of human origin were retrieved from each biological replicate (96-250×10^6/each) as compared to the 3,544 peptides retrieved using equivalent or lower total numbers (100×10^6) of untreated β cells in our previous work (15). Besides causing β-cell lysis and self-antigen release, the second pathogenic effect of CVB infection is the triggering of type I IFN release, which creates an immunogenic environment and induces the secondary HLA-I hyper-expression found to co-localize with VP1 staining in islets of T1D patients (10) and persisting years after clinical diagnosis (32). This HLA-I hyperexpression leads to an increased peptide presentation (15), thus enhancing the autoimmune vulnerability of β cells.

This study carries limitations. First, only the peptides naturally processed and presented by β cells were considered. While this allowed us to directly address the possibility that vaccination may trigger $CD8^+$ T-cell cytotoxicity against β cells, the peptide display offered for T-cell activation by other cells, including antigen-presenting cells, may be broader, as suggested by

the identification of additional epitopes by *in-silico* predictions. Notably, it will be important to define the peptides displayed by infected enterocytes at the intestinal viral entry site, and by dendritic cells and macrophages that may take up antigens locally and in draining lymph nodes. Second, the contribution of $CD4^+$ T-cell responses to CVB immunity remains to be explored. Third, whether $CD8^+$ T-cell responses are weaker in children subsequently developing islet autoimmunity remains an open question.

In conclusion, the identification of the CVB epitopes targeted by $CD8^+$ T cells allowed us to define the major role of CVB infection in mediating direct β-cell lysis rather than relying on anti-viral $CD8^+$ T-cell cytotoxicity, which mitigates the risk that CVB vaccination may precipitate the destruction of infected β cells. The poor $CD8^+$ T-cell memory against these CVB epitopes lends rationale for boosting such memory through vaccination.

ACKNOWLEDGMENTS. We thank all study participants for providing blood samples and the Cochin Institute CyBio Flow Cytometry Core Facility for assistance with cell analysis and sorting. We gratefully acknowledge Dr. Vesa Hytönen (University of Tampere, Finland) for the gift of the anti-VP1 mAb 3A6; and Dr. Lindsay Whitton (The Scripps Research Institute, La Jolla, CA, USA) for the gift of the CVB-GFP strain.

FUNDING. This work was supported by JDRF nPOD-Virus grant 3-SRA-2017-492-A-N (to A.P.), and Postdoctoral Fellowship 3-PDF-2020-942-A-N (to Z.Z.); the *Agence Nationale de la Recherche* (ANR-19-CE15-0014-01) and the *Fondation pour la Recherche Medicale* (EQU20193007831), to R.M. S.T. was supported by the Novo Nordisk Postdoctoral Fellowship program at Karolinska Institutet. Work in the laboratory of T.R.C., R.S., M.S. and R.M. received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, the European Federation of Pharmaceutical Industries and Associations, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust. Work in the laboratory of M. F-T. was supported by the Strategic Research Program in Diabetes at Karolinska Institutet and the Swedish Research Council. This research was performed with the support of the Network for Pancreatic Organ donors with Diabetes (nPOD; RRID: SCR_014641), a collaborative type 1 diabetes research project sponsored by JDRF (nPOD: 5-SRA-2018-557-Q-R) and The Leona M. & Harry B. Helmsley Charitable Trust (Grant #2018PG-T1D053). Organ Procurement Organizations (OPO) partnering with nPOD to provide research resources are listed at www.jdrfnpod.org/for-partners/npod-partners. The content and views expressed are the responsibility of the authors and do not necessarily reflect an official view of nPOD.

DUALITY OF INTEREST. R.M. receives research funding from Provention Bio Inc. M.F.–T. serves on the scientific advisory board of Provention Bio Inc.

METHODS

Preparation and titration of CVB1 and CVB3 stocks

CVB1 (ATCC strain), CVB3 (Nancy strain) and GFP-CVB3 (33) were grown in HeLa cells. Supernatants and cells were collected after overnight culture, subjected to 3 freeze/thaw cycles and centrifuged at 4,000 rpm. CVB1 and CVB3 were further purified by ultracentrifugation. Next-generation sequencing was used to confirm serotypes and identify sequence variants (quasi-species). Virus stocks were titrated on HeLa cells.

In-vitro infection

The ECN90 β-cell line (HLA-A*02:01/03:01, HLA-B*40:01/49:01, HLA-C*03:04/07:01) (16) was cultured in DMEM/F12 Advanced medium (ThermoFisher) supplemented with 2% bovine serum albumin (BSA) fraction V, fatty-acid-free (Roche), 50 μM β-mercaptoethanol (Sigma), 10mM nicotinamide (Merck), 1.7 ng/mL sodium selenite (Sigma), 100U/mL Penicillin-Streptomycin (Gibco), 1% glutaMAX (Gibco). Cells were seeded at 16.5 x 10^6 cells in 150 cm^2 culture flasks (TPP) coated with 0.25% fibronectin from human plasma (Sigma) and 1% extracellular matrix from Engelbreth-Holm-Swarm murine sarcoma (Sigma) and cultured at 37°C in 5% CO_2 for 16-24h. They were infected with 300 MOI of CVB1 or CVB3 for 1 h in BSA-free supplemented medium. After gentle washing, the culture was pursued for 6 h, followed by trypsinization, washing with phosphate-buffered saline and processing for flow cytometry or freezing at –80°C as dry pellets for peptidomics experiments.

Infection was verified on cells stained with Live/Dead Violet (ThermoFisher) and monoclonal antibodies (mAbs) to HLA-A,B,C (RRID:AB_2917755), VP1 (clone 3A6) (34), dsRNA (RRID:AB_2651015) followed by a FITC-conjugated mouse anti-rat IgG (RRID:AB_465229) and an AlexaFluor (AF)594-conjugated goat anti-mouse IgG (RRID: AB_2762825). Western blot for HLA-A,B,C was performed on whole cell lysates using the mAb HC10 (RRID: AB_2728622; produced in-house) and a mAb to α-tubulin (RRID: AB_1210457) as loading control, followed by a horseradish peroxidase-conjugated goat anti-mouse IgG (RRID: AB_2728714) and Pierce ECL reagents (ThermoFisher).

Isolation and identification of HLA-I-bound peptides

Frozen cell pellets were lysed for 1 h at 4°C with 1% (w/v) octyl-b-D glucopyranoside (Sigma) diluted in 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA and 0.1% (v/v) Complete Protease Inhibitor Cocktail (Roche), with vortexing every 30 min. Lysates were cleared by

centrifugation and pHLA-I complexes isolated by immunoaffinity purification using W6/32 mAb (produced in-house) crosslinked to protein A beads (GE Healthcare) with dimethyl pimelimidate (Sigma). Beads were loaded into GELoader Tips (ThermoFisher), washed and pHLA-I complexes eluted with 10% (v/v) acetic acid. The eluates were further concentrated by SpeedVac, acidified with 10% trifluoroacetic acid (TFA, ThermoFisher) and loaded into pre-washed and -equilibrated C18 stage tips (ThermoFisher). After further washing, peptides were eluted with 50% (v/v) acetonitrile (ACN, ThermoFisher) and 0.05% (v/v) TFA. Samples were dried, resuspended in 0.05% TFA 2% ACN and spiked with 10 fmol/µL of CMV $pp65_{495-503}$ peptide (NLVPMVATV) as an internal control.

Using a Dionex Ultimate 3000 RSLCnano system, peptides were loaded onto a C18 Acclaim PepMap nanocolumn (5 mm x 300 µm; ThermoFisher) and separated with an Acclaim PepMap C18 nanocolumn (50 cm x 75 µm; Dionex) coupled to a nano-electrospray ionization Q Exactive HF mass spectrometer (ThermoFisher). Separation was achieved with a linear gradient of 2-90% buffer B (80% ACN, 0.05% formic acid) at a flow rate of 220 nL/min over 90 min at 35°C.

Full MS spectra were acquired from 375 to 1,500 m/z, at a resolution of 60,000 with an automatic gain control (AGC) target of 3,000,000. Precursors were selected using Top20 mode at a resolution of 15,000 with an AGC target of 50,000, a maximum injection time of 120 ms and a dynamic exclusion of 20 s. Unassigned and >4 precursor ion charge states were excluded. The peptide match option was set to 'preferred'.

Data was analyzed using Peaks 8.1 (Bioinformatics Solution). Sequence interpretation was carried out using no enzyme specification, a mass tolerance precursor of 10 ppm for peptides and 0.05 for fragment ions. Sequence matching was done against a database containing: a) all human SwissProt entries (downloaded 11/2017); b) translated sequences of RNA-sequenced viral strains; c) an in-house database compiled as described (15), containing 679,576 and 530,424 predicted peptide splice products composed of major known and candidate β-cell antigen sequences with CVB1 and CVB3 sequences, respectively; d) large T antigen from simian virus (UniProt P03070) used for β-cell immortalization; and e) putative polyproteins from a xenotropic murine leukemia virus (UniProt F8UU35, F8UU37) known to be present in ECN90 cells (35). Post-translational modifications detected with the Peaks PTM built-in module were disregarded as likely artifactual and the unmodified aa sequence was retained. A false discovery rate of 5% was set using a parallel decoy database search. Only the peptides with a 8-12 aa length compatible with HLA-I binding were retained.

In-silico selection of candidate epitopes

A parallel *in-silico* search for nonamer peptides predicted to bind HLA-A2 or HLA-A3 was performed for all 6 CVB strains, using the following reference sequences: CVB1 AAC00531.1 (UniProt P08291), CVB2 AOW42548.1 (UniProt A0A1D8QMF3), CVB3 AFC88096.1 (UniProt H9B4F8), CVB4 AHB37371.1 (UniProt W8DN51), CVB5 AFO42818.1 (UniProt I7AVS5), CVB6 AAF12719.1 (UniProt Q9QL88). The aa positions of each peptide are reported with reference to the first serotype listed. Criteria used for this search were: a binding affinity (IC50) <40 nM, as predicted by both NetMHCpan 4.0 through the IEDB HLA-I prediction interface and NetMHC 4.0; or an identity of ≥3 consecutive aa with a β-cell antigen sequence (INS, GAD, IA-2, ZnT8, IAPP, IGRP) in central TCR contact residues (positions P3-P7). Peptides containing cysteines or found only in CVB2 or CVB6 were excluded. Two CVB peptides previously reported (36) were appended to this list: CVB1-2-3-4-5-$6_{1502-1510}$ (GIIYIIYKL, IEDB #20311) and CVB2-4-5-$6_{1589-1598}$ (ILMNDQEVGV, IEDB #27205).

HLA-I binding assays

Peptides predicted to bind HLA-A2 or -A3 (NetMHCStabPan 1.0) were synthesized (>95% pure, Synpeptide) and binding tested using biotin-tagged HLA-A2 and -A3 monomers (ImmunAware), as described (17). Briefly, biotinylated monomers were folded (1.2 nM final concentration) and captured on streptavidin-coated beads (6-8µm, Spherotech), followed by incubation with anti-β2-microglobulin BBM.1 monoclonal antibody (Santa Cruz) and then with AF488-labeled goat polyclonal anti-mouse IgG (RRID:AB_2728715). Positive controls included the HLA-A2-binding peptides Flu MP_{58-66} (GILGFVFTL) and $TYR_{369-377}$ (YMDGTMSQV) and the HLA-A3-binding peptide Flu $NP_{265-273}$ (ILRGSVAHK). The peptide $CHGA_{382-390}$ (HPVGEADYF) was used as negative control for both HLA-A2 and -A3. Beads were acquired on a BD LSRFortessa cytometer and analyzed by gating on single beads and AF488-positive events. Results are expressed as relative fluorescence intensities (RFIs), i.e. the median fluorescence intensity fold increase of the tested pHLA-I complex compared with the negative control. For HLA-A2, peptides with RFI>6 were considered as intermediate binders and RFI>20 as strong binders. For HLA-A3, peptides with RFI>15 were scored as intermediate binders and RFI>40 as strong binders.

Study participants, specimens, HLA typing and CVB serology

Study participants were recruited in Paris and Miami and gave written informed consent under ethics approval DC-2015-2536 Ile-de-France I (Paris) and University of Miami IRB protocol 1995-119 (Miami). HLA-A2 (A*02:01) and -A3 (A*03:01) typing was performed with AmbiSolv primers (Dynal/ThermoFisher) and by DKMS (Tübingen, Germany) in case of typing ambiguities. Blood was drawn into 9-mL sodium heparin tubes and PBMCs isolated and frozen/thawed as described (15, 37). Frozen/thawed samples were used throughout the study. Neutralizing antibody titers against the 6 CVB serotypes were measured with a plaque neutralization assay as described (38). Frozen PLN, spleen and blood cells were provided by the Network for Pancreatic Organ donors with Diabetes (nPOD) and thawed in the presence of benzonase (50 U/ml, Merck 70664-3) and mifepristone (100 nM, Invitrogen H110-01).

Combinatorial HLA Class I multimer assays

HLA-A2 and -A3 multimers (MMrs) (immunAware) were produced and used as described (15, 17, 18). pHLA complexes were conjugated with fluorochrome-labeled streptavidins (1:4 ratio) and used at a final concentration of 8-27 nmol/L, modulated in order to correct for the variable staining index of each streptavidin and thus visualize a distinct double-MMr^+ population for each fluorochrome pair. PBMCs were thawed in prewarmed AIM-V medium (ThermoFisher), washed, counted and incubated with 50 nmol/L dasatinib for 30 min at 37°C before magnetic depletion of $CD8^-$ cells (RRID:AB_2728716). Staining with combinatorial double-coded MMr panels was performed for 20 min at 20°C in 20 µL PBS-dasatinib for 10^7 cells followed, without washing, by staining at 4°C for 20 min with monoclonal antibodies CD3-APC-H7 (RRID:AB_1645475), CD8-PE-Cy7 (RRID:AB_396852), CD45RA-FITC (RRID:AB_395879), CCR7-BV421 (RRID:AB_2728119), CD25-BV510 (RRID:AB_2629671), PD-1-BV605 (RRID:AB_2563212) and Live/Dead Aqua (ThermoFisher). After washing, cells were acquired on an LSRFortessa configured as described (15) and analyzed with FlowJo v10 and GraphPad Prism 8. Candidate epitopes that did not yield any appreciable MMr staining or fluorochrome pairs not coding for a pHLA MMr provided negative controls for each panel. Positive control peptides included in the MMr panels were HCV $PP_{1406-1415}$ (naïve viral control; KLVALGINAV; IEDB # 32208) and/or $MelanA_{26-35}$ ELA (naïve self control; ELAGIGILTV; IEDB #12941) and Influenzavirus (Flu) MP_{58-66} (recall viral control; GILGFVFTL; IEDB #20354) for HLA-A2; and HIV Nef_{84-92} (naïve viral control; AVDLSHFLK; IEDB #5295) and Influenzavirus $NP_{265-273}$ (recall viral control; ILRGSVAHK; IEDB #27283) for HLA-A3.

For the detection of $CVB1_{1132-1140}/GAD_{272-280}$ cross-reactive $CD8^+$ T cells, the $CVB1_{1132-1140}$ MMr pair (PE-CF594 and BV711) shared the BV711 fluorochrome with the $GAD_{272-280}$ MMr pair (PE and BV711), thus allowing selective gating of T cells recognizing only $CVB1_{1132-1140}$ or $GAD_{272-280}$ (i.e., double MMr^+) or both (i.e., triple MMr^+), as described (18).

In-vitro T-cell stimulation, TCR sequencing and re-expression

To generate PLN T-cell lines, frozen/thawed single-cell suspensions were bulk-sorted by flow cytometry for $CD4^+$ and $CD8^+$ T cells. $CD8^+$ T cells (10^5/well, 4 wells/each) were stimulated with anti-CD28 (5 µg/ml) and irradiated autologous spleen-derived EBV-immortalized B cells pulsed with the indicated peptides (50µg/ml), with IL-2 (20 U/ml) and IL-15 (10 ng/ml) added on day 6. On day 14, cultures were split and recalled with 50 µg/ml of the stimulating peptide, the negative control peptide West Nile Virus $[WNV]_{3098-3106}$ or DMSO diluent pulsed on irradiated autologous B cells with anti-CD28. After 48 h, supernatants were collected and assayed for cytokine secretion by cytokine bead array (BD).

Single-cell TCR sequencing was performed on frozen/thawed cells rested overnight, followed by $CD8^+$ magnetic enrichment and MMr staining, sorting in 96-well plates with a BD FACSAria II and a multiplex nested PCR, as described (16). Sequences were assigned using IMGT (https://www.imgt.org) and visualized as heatmaps using R software (https://github.com/raivokolde/pheatmap).

5KC T-hybridoma TCR transductants were generated as described (23). Briefly, 5KC cells were transduced with the NFAT-driven fluorescent reporter ZsGreen-1 along with human CD8 by spinoculation with retroviral supernatant produced from Phoenix-ECO cells (ATCC CRL-3214). These cells were subsequently transduced with retroviral vectors encoding a murine TRAC chimeric TCRα gene followed by a porcine Teschovirus-1 2A (P2A) peptide and a murine TRBC chimeric TCRβ gene (Twist Bioscience.

Primary $CD8^+$ T-cell transductants were generated from $CD8^+$ T cells magnetically enriched from fresh PBMCs and plated at 20,000 cells/well. After activation with anti-CD3 and anti-CD28 microbeads (ThermoFisher) for 48 h in AIM-V medium, cells were transduced with lentiviral vectors encoding the above chimeric TCR construct and supplemented with 40 U/ml IL-2 (Proleukin, Clinigen).

To validate antigen specificity, transductants (20,000 cells/well in 96-well U-bottom plates) were stimulated for 18 h with individual peptides added at 10-fold serial dilutions from 100 µM to 100 fM in the presence of K562 cells (50,000 cells/well) transduced with an HLA-A2 or HLA-A3 monochain construct fused with β2-microglobulin, followed by analysis of ZsGreen-

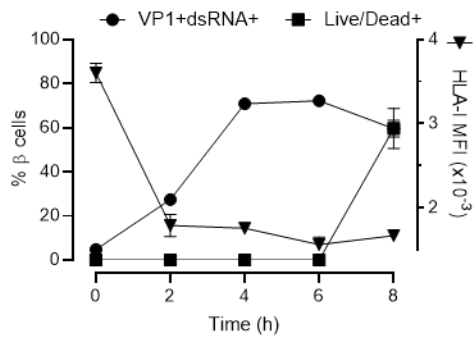
1 expression (for 5KC transductants) or intracellular staining for MIP-1β (RRID:AB_357303), TNF-α (RRID:AB_398566), IL-2 (RRID:AB_2573518) and IFN-γ (RRID:AB_1272026) for primary $CD8^+$ transductants.

In-vitro recall and cytotoxicity assays on CVB-infected β cells

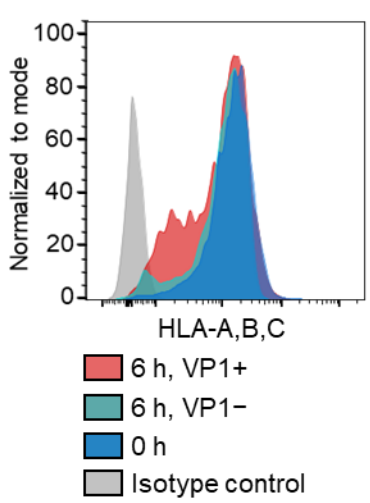
For recall assays of 5KC TCR transductants, the HLA-$A3^+$ EndoC-βH2 β-cell line was plated 24 h ahead at $5x10^4$/well in a 96-well flat-bottom Matrigel-coated plate in in Advanced DMEM/F12 medium (ThermoFisher) supplemented with GlutaMax, 1% penicillin/streptomycin, 50 µM 2-mercaptoethanol, 10 mM nicotinamide and 3.4 ng/L sodium selenite. It was then pulsed for 1 h with serial concentrations of $CVB1_{1356-1364}$/$CVB3_{1359-1367}$ peptide or left unpulsed, followed by CVB3 infection (300 MOI) for an additional hour, gentle washing, and addition of transductants. After 18 h of coculture, cells were collected, stained with Live/Dead Red and fixed with the eBioscience Foxp3/Transcription Factor Staining Buffer Set. NFAT-driven ZsGreen fluorescence was then acquired on a BD LSRFortessa after gating on viable single cells.

For cytotoxicity assays, primary $CD8^+$ T-cell transductants were co-cultured for 12 h at different effector:target ratios with ECN90 target cells seeded at 40,000 cells/well in pre-coated 96-well plates. Target cells were preliminarily loaded with 1 µM CVB peptides or infected with GFP-CVB at different MOI for 1 h at 37°C. The percent specific lysis was defined as the number of remaining viable cells/µl acquired for each condition as compared to control wells (uninfected cells).

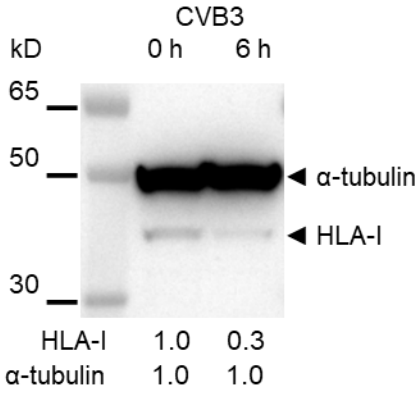
A



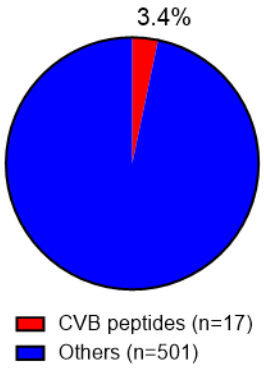
B



C



D



E

Strain	Position	Sequence	Number of hits	Predicted restriction	Predicted affinity (nM) / stability (h)	*In vitro HLA* binding	T-cell screening	T-cell validation
CVB3	159-167	GLFGQNMQY	2/2	A*03:01	142 / 0.46	–	ND	ND
CVB3	236-244	LVQRVVYNA	1/2	A*02:01	1852 / 1.10	ND	ND	ND
CVB3	*235-244*	*KLVQRVVYNA*	*NA*	*A*02:01*	*56 / 5.75*	++	–	ND
CVB1	804-812	RIYFKPKHV	1/1	NA	NA	ND	ND	ND
CVB1	*804-813*	*RIYFKPKHVK*	*NA*	*A*03:01*	*16 / 5.56*	++	+	–
ZnT8	237-246	IIYFKPEYKI	NA	A*03:01	8769 / 0.15	+	ND	ND
CVB3	1038-1045	SILEKSLK	2/2	A*03:01	98 / 1.24	–	ND	ND
CVB3	*1038-1048*	*SILEKSLKALV*	*NA*	*A*02:01*	*89 / 1.52*	+	–	ND
CVB3	1092-1100	YGIPMAERQ	1/2	NA	NA	ND	ND	ND
CVB1	1132-1140	KILPEVKEK	1/1	A*03:01	504 / 1.79	++	+	+
CVB3	1135-1143	KILPEVREK	2/2	A*03:01	317 / 2.29	++	ND	+
GAD	272-280	KMFPEVKEK	NA	A*03:01	48 / 4.70	++	ND	+
CVB3	1203-1211	SLEKKMSNY	1/2	A*03:01	3016 / 0.18	–	ND	ND
CVB1	1233-1241	SVATNLIGR	1/1	A*03:01	1283 / 0.39	++	–	ND
CVB3	*1236-1244*		*2/2*					
CVB3	1339-1347	ARRFHFDMN	1/2	NA	NA	ND	ND	ND
CVB1	1356-1364	KINMPMSVK	1/1	A*03:01	17 / 3.84	++	+	+
CVB3	1359-1367		1/2					
CVB1	*1356-1363*	*KINMPMSV*	*NA*	A*02:01	386 / 1.62	+	–	ND
CVB3	*1359-1366*		*NA*					
CVB3	1395-1403	RRTQVRYSL	2/2	C*07:01	209 / 0.12	ND	ND	ND
CVB3	1519-1529	GAYTGVPNQKP	1/2	NA	NA	ND	ND	ND
CVB3	*1519-1528*	*GAYTGVPNQK*	*NA*	*A*03:01*	*123 / 1.21*	++	–	ND
CVB1	1536-1543	VQGPVFEF	1/1	NA	7243 / 0.48	ND	ND	ND
CVB1	*1536-1545*	*KVQGPVFEFAV*	NA	*A*02:01*	*91 / 1.99*	–	ND	ND
CVB3	1541-1549	GPAFEFAVA	1/2	NA	NA	ND	ND	ND
CVB1	1557-1565	TEYGEFTML	1/1	B*40:01	11 / 1.51	ND	ND	ND
CVB3	1560-1568		2/2					
CVB3	1765-1774	VLRSGDPRLK	2/2	A*03:01	152 / 0.99	++	+	+
CVB3	1924-1934	NPGVVTGSAVG	1/2	NA	NA	ND	ND	ND

F

Serotype	Position	Sequence	Predicted restriction	Predicted affinity (nM) / stability (h)	*In vitro HLA* binding	T-cell screening	T-cell validation
CVB1-2-5	60-68	IMIKSMPAL	A*02:01	21 / 3.43	++	+	–
CVB4-6	60-68	VMIKSLPAL	A*02:01	16 / 5.89	++	–	ND
CVB5	169-177	YLGRAGYTV	A*02:01	12 / 3.30	++	+	+
CVB1-2-4-5-6	271-279	IVMPYINSV	A*02:01	6 / 12.48	++	+	+
CVB2	291-299	FTLMIIPFV	A*02:01	7 / 2.71	++	–	ND
CVB1-2-3-4-5-6	456-464	AMATGKFLL	A*02:01	36 / 1.98	++	–	ND
CVB1-5	555-563	RMLKDTPFI	A*02:01	14 / 3.71	++	–	ND
CVB2-5	711-719	VLTHQIMYV	A*02:01	7 / 21.90	+	–	ND
CVB5	760-768	SMFYDGWAK	A*03:01	27 / 0.97	+	+	+
CVB4	943-951	YQSHVLLAV	A*02:01	9 / 14.43	++	–	ND
CVB5	945-953	YQTHVLLAV	A*02:01	21 / 13.04	++	+	+
CVB1-2-3-4-5-6	1204-1212	KMSNYIQFK	A*03:01	8 / 8.93	++	–	ND
CVB1-2-3-4-5-6	1246-1254	KLNSSVYSL	A*02:01	11 / 5.52	++	+	+
CVB1-2-3-4-5-6	1289-1297	QMVSSVDFV	A*02:01	18 / 2.45	++	–	ND
CVB1-2-3-4-6	1307-1315	GILFTSPFV	A*02:01	28 / 1.66	++	–	ND
CVB1-2-3-4-6	1308-1316	ILFTSPFVL	A*02:01	19 / 2.61	+	ND	ND
CVB5	1311-1319	FLFTSPFVL	A*02:01	4 / 5.06	++	–	ND
CVB2	1361-1369	KVNMPMSVK	A*03:01	20 / 5.66	++	+	ND
CVB1-2-3-4-5-6	1509-1517	KLFAGFQGA	A*02:01	13 / 4.82	+	–	ND
CVB1-3	1585-1593	LMNDQEIGV	A*02:01	14 / 2.59	++	–	ND
CVB2-4-5-6	1589-1598	ILMNDQEVGV	A*02:01	9 / 9.86	++	+	–
CVB1-5-6	1626-1634	FLAKEEVEV	A*02:01	8 / 5.09	++	–	ND
CVB2-4	1631-1639	FLAREEAEV	A*02:01	11 / 3.95	+	–	ND
IA-2	671-679	RLAKEWQAL	A*02:01	312 / 1.11	++	ND	ND
CVB1-2-3-4-5-6	1671-1679	RMLMYNFPT	A*02:01	20 / 3.12	+	+	+
CVB1-2-3-4-5-6	2024-2032	IIIRTLMLK	A*03:01	22 / 1.87	++	–	ND
CVB1-2-3-4-5-6	2027-2035	RTLMLKVYK	A*03:01	43 / 6.37	++	–	ND
CVB1-2-3-4-5-6	2107-2115	FLVHPVMPM	A*02:01	20 / 2.41	++	–	ND
CVB2-3-4-5	2161-2169	KIRSVPVGR	A*03:01	469 / 0.85	++	–	ND
IGRP	232-240	LLWSVPIAK	A*03:01	35 / 1.64	++	ND	ND

Figure 1. CVB infection of ECN90 β cells and candidate epitopes identified. A. ECN90 β cells were infected *in vitro* with CVB3 at 300 MOI. The kinetics of expression of the viral markers VP1 and dsRNA and of cell death (Live/Dead$^+$ cells) was measured by flow cytometry. **B-C.** HLA-I expression measured by flow cytometry (B) and Western blot (C) on ECN90 β cells 6 h post-infection. **D.** Percent CVB peptides out of total 8-12mer peptides eluted. **E.** CVB peptides retrieved from HLA peptidomics experiments and length variants of the core sequences identified (shown in italics). The number of hits out of the 3 peptidomics runs performed (2 for CVB3 and 1 for CVB1) are listed for each peptide. A $ZnT8_{237-246}$ peptide homologous to $CVB1_{804-812}$ (identified by Blast search) and a $GAD_{272-280}$ peptide homologous to the $CVB1_{1132-1140}$/$CVB3_{1135-1143}$ sequence retrieved from our previous datasets (15) are also listed. HLA-I restrictions were assigned using NetMHCStabpan 1.0, and predicted affinity and stability values are listed. Subsequent columns list *in-vitro* HLA binding results (detailed in Fig. S3) and peptides identified as recognized by CD8$^+$ T cells in the screening (Fig. 2) and validation runs (Fig. 3). Peptides eventually validated for T-cell recognition are shaded in grey. $CVB1_{1132-1140}$ and $CVB3_{1135-1143}$ differ only by 1 aa and were thus counted as one peptide (total n=17). **F.** CVB nonamer peptides identified in a parallel *in-silico* search for all 6 CVB strains. Homologous β-cell peptide sequences identified by Blast search are also listed. Three peptides that did not confirm as binders *in vitro* were excluded and are detailed in Fig. S3. For HLA-eluted peptides, sequence annotations refer to the CVB1/CVB3 strains used to infect β cells. For *in-silico* predicted peptides, annotations indicate sequence identities across serotypes. NA, not applicable or not assigned; ND, not determined.

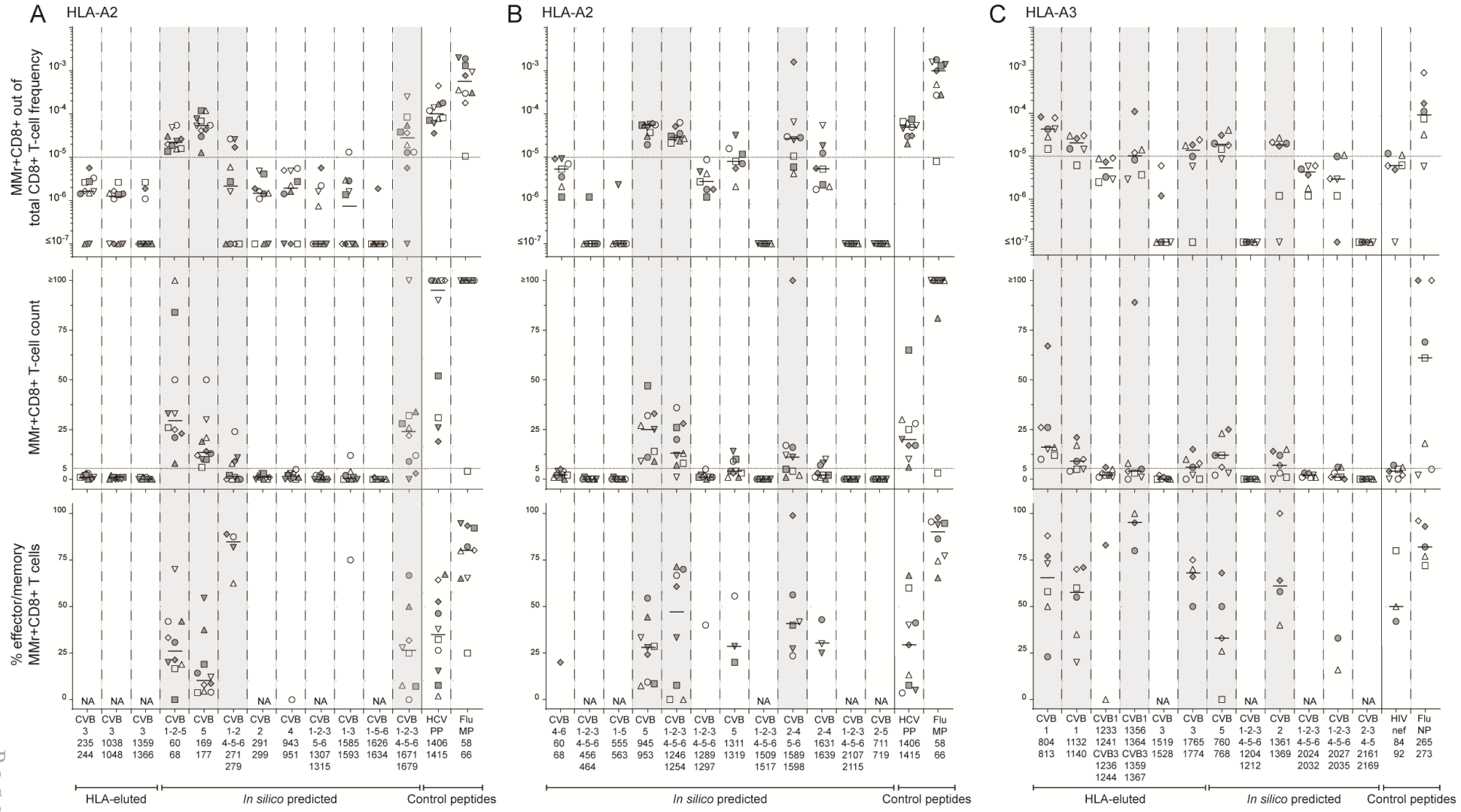

A
HLA-A2
B
HLA-A2
C
HLA-A3
MMr+CD8+ out of total CD8+ T-cell frequency
MMr+CD8+ T-cell count
% effector/memory MMr+CD8+ T cells
HLA-eluted
In silico predicted
Control peptides

Figure 2. Screening of CVB peptides for recognition by blood CD8$^+$ T cells in seropositive healthy adults. A-C. HLA-A2-rectricted (A-B) and HLA-A3-restricted candidates (C) were tested with combinatorial MMr panels. Each symbol represents a donor (legends in Table S1) and the bars display the median value of each distribution. For each panel, the top graph depicts the frequency of MMr$^+$CD8$^+$ T cells out of total CD8$^+$ T cells, with the horizontal line indicating the 10^{-5} frequency cut-off used as a first validation criterion; the middle graph displays the number of MMr$^+$CD8$^+$ T cells counted, with the horizontal line indicating the cut-off of 5 cells used to assign an effector/memory phenotype; the bottom graph shows the percent fraction of effector/memory cells (i.e. excluding naïve CD45RA$^+$CCR7$^+$ cells) among MMr$^+$CD8$^+$ T cells (for those donors with ≥5 cells counted; NA when not assigned). Peptides validated in this screening phase are highlighted in grey. In each panel, control peptides are derived from viruses eliciting predominantly naïve responses in these unexposed individuals (HCV $PP_{1406-1415}$ and HIV nef_{84-92} for HLA-A2 and HLA-A3, respectively) and from Influenzavirus, eliciting predominantly effector/memory responses (Flu MP_{58-66} and Flu $NP_{265-273}$ for HLA-A2 and HLA-A3, respectively).

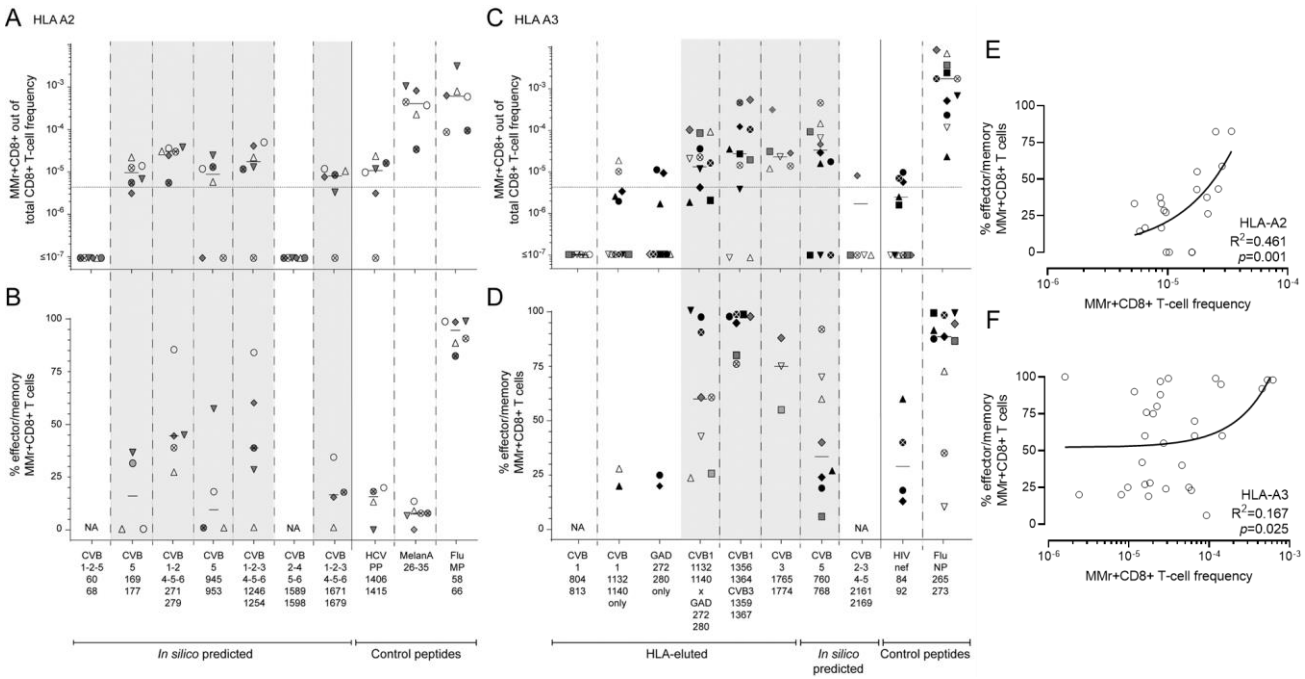


Figure 3. Validation of CVB candidate epitopes as targets of blood CD8$^+$ T cells in seropositive healthy adults. A-D. The HLA-A2-rectricted and HLA-A3-restricted peptides selected for these validation experiments from the screening round (Fig. 2) are shown in panels A-B and C-D, respectively. In the HLA-A3 panel, the CVB2-3-4-5$_{2161-2169}$ peptide not retained after screening was here included as negative control. Each symbol represents a donor (legends in Table S1), with donors recruited in Paris and Miami depicted in panels C-D as white/gray and black symbols, respectively. The bars display the median values of each distribution. Panels A, C depict the frequency of MMr$^+$CD8$^+$ T cells out of total CD8$^+$ T cells, with the horizontal line indicating the $5x10^{-6}$ frequency cut-off used for validation, with only peptides scoring ≥3 MMr$^+$ cells retained for this more stringent analysis. Panels B, D show the percent fraction of effector/memory cells among MMr$^+$CD8$^+$ T cells. Validated peptides are highlighted in grey. In each panel, control peptides are derived from viruses or self-antigens eliciting predominantly naïve responses (HCV PP$_{1406-1415}$/MelanA$_{26-35}$ and HIV nef$_{84-92}$ for HLA-A2 and HLA-A3, respectively) and from Influenzavirus, eliciting predominantly effector/memory responses (Flu MP$_{58-66}$ and Flu NP$_{265-273}$ for HLA-A2 and -A3, respectively). For the CVB1$_{1132-1140}$ peptide homologous to GAD$_{272-280}$, frequencies and effector/memory fractions are shown for T cells recognizing either peptide or both, as detailed in Fig. S6. **E-F.** Correlation between the frequency of CVB epitope-reactive MMr$^+$CD8$^+$ T cells and their percent effector/memory fraction for HLA-A2- (E) and HLA-A3-restricted peptides (F). Each symbol represents an individual epitope specificity in an individual donor

.

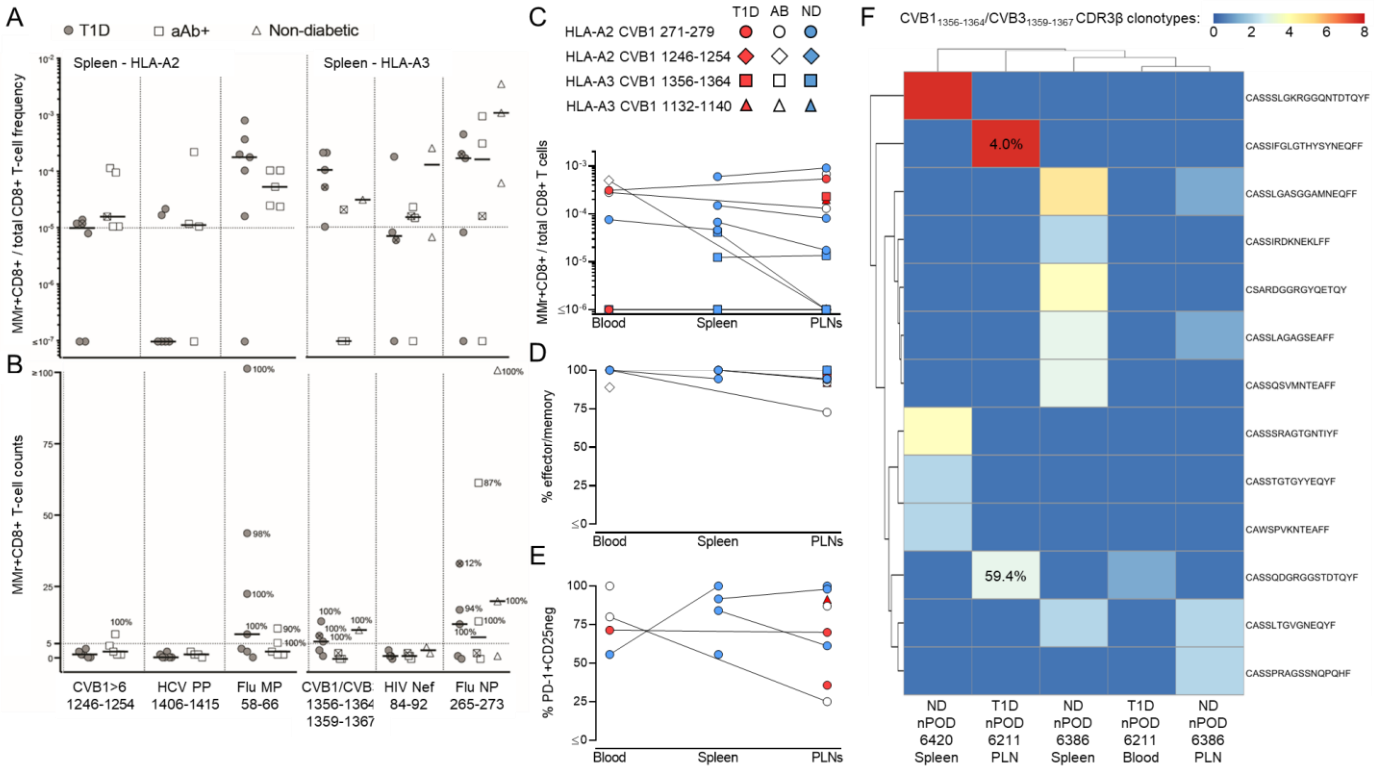


Figure 4. Detection of CVB-reactive CD8$^+$ T cells in lymphoid tissues of nPOD donors. A-B. Detection of CVB epitope-reactive CD8$^+$ T cells in splenocytes from nPOD cases (HLA-A2$^+$, left; and HLA-A3$^+$, right; see details in Table S4). Each symbol represents a donor group (with 2 HLA-A2/A3$^+$ donors indicated by crossed symbols) and the bars indicate the median values of each distribution. Panel A depicts the frequency of MMr$^+$CD8$^+$ T cells out of total CD8$^+$ T cells, with the horizontal line indicating the 10^{-5} frequency cut-off. A median of 73,320 CD8$^+$ T cells were counted (range 9,611-307,697). Panel B displays the number of MMr$^+$CD8$^+$ T cells counted, with the horizontal line indicating the cut-off of 5 cells used to assign an effector/memory phenotype. For those donors/epitopes with ≥5 MMr$^+$ cells counted, the percent effector/memory fraction within each MMr$^+$CD8$^+$ population is annotated. Only epitopes testing positive are depicted (complete peptide panels listed in Table S5). **C-E.** CVB epitope-reactive CD8$^+$ T cells in lymphoid tissues from nPOD cases (see details in Table S4). MMr$^+$CD8$^+$ T-cell frequencies, percent effector/memory MMr$^+$ and PD-1$^+$CD25$^-$ effector/memory MMr$^+$ cells (for donors/epitopes with cell counts ≥5) are depicted across available tissues in the top, middle and bottom panel, respectively. **F.** Expanded TCR CDR3β clonotypes in individual CBV1/CVB3$_{1356\text{-}1364/1359\text{-}1367}$-reactive CD8$^+$ T cells from the same nPOD cases/tissues. Expanded clonotypes were defined as those found in at least 2 MMr$^+$ cells in the same tissue from the same donor, and are clustered according to frequency (one column per tissue/case) and sequence similarity (one row per CDR3β clonotype). For the PLN CD8$^+$ T cells of T1D nPOD case 6211 (second column), percentages indicate the CDR3β clonotype frequency out of total PLN CD8$^+$ T cells sequenced by Seay et al. (21), available at http://clonesearch.jdrfnpod.org/.

A

CVB epitope	HLA restriction	CDR3β / CDR3α
CVB1 1356-1364	HLA-A3	CASSLKGNNEQFF CAVGTTNAGKSTF

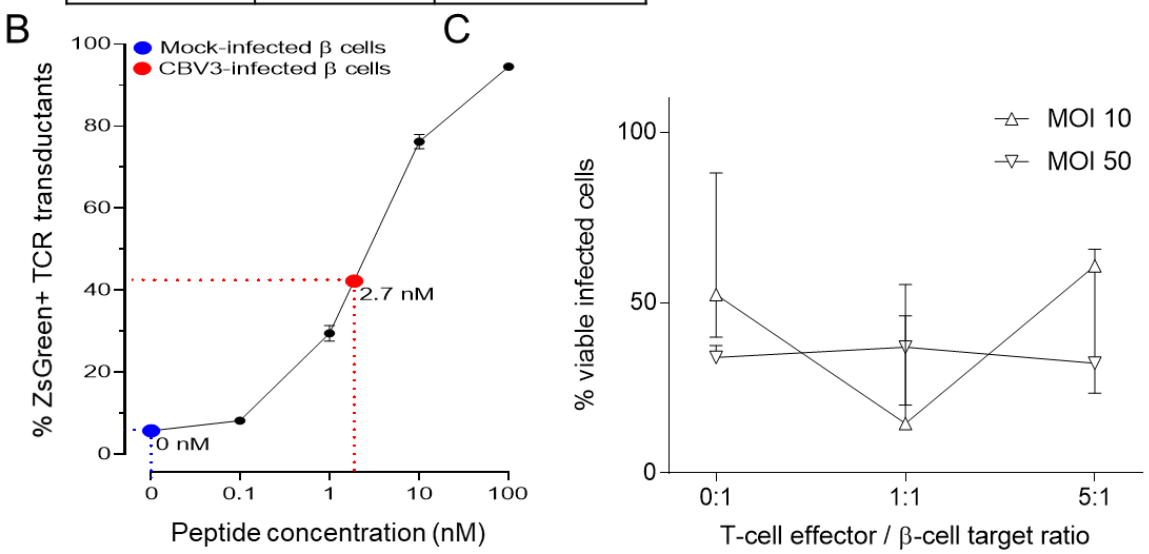


Figure 5. Lysis of infected β cells by CVB and CVB-reactive CD8$^+$ T cells. A. HLA-A3-restricted CVB1$_{1356\text{-}1364}$-reactive TCR sequenced and re-expressed in carrier cells. **B.** T-cell activation in NFAT-driven ZsGreen fluorescent reporter TCR-transduced 5KC cells upon an 18-h co-culture with CVB3-infected (red; 300 MOI) vs. mock-infected β cells (blue). A parallel dose-response curve with increasing peptide concentrations pulsed on non-infected β cells (black) estimated the concentration of peptide recognized on CVB-infected β cells to 2.7 nM. **C.** β-cell lysis at increasing CVB3 MOI and increasing T-cell effector to target ratios, using primary CD8$^+$ T cells transduced with the same TCR. Data is depicted as median ± range of triplicate wells.

SUPPLEMENTAL MATERIAL

Table S1. CVB-seropositive healthy donors recruited for T-cell screening and validation experiments. Columns display donors' demographics, HLA-A2/A3 typing, titers of neutralizing antibodies (NAbs) for different CVB serotypes (positive titers of ≥16 in red, borderline titers of 4 in blue; corresponding to serum dilutions of 1/16 and 1/4, respectively), inclusion into T-cell screening and/or validation experiments and the corresponding symbols used in Fig. 2 and 3, with HLA-A2^{+} donors recruited in Miami indicated with black symbols. All other donors were recruited in Paris.

Code	Age (yrs)	Gender (M/F)	HLA-A2	HLA-A3	CVB1 NAbs	CVB2 NAbs	CVB3 NAbs	CVB4 NAbs	CVB5 NAbs	CVB6 NAbs	Screening round	Validation round	Symbols
H110S	60	F	N	Y	0	≥16	0	≥16	≥16	4	+		
H112S	42	F	N	Y	≥16	≥16	0	≥16	4	≥6	+		
H423C	26	M	N	Y	na	na	na	na	na	na		+	
H458C	33	M	N	Y	≥16	≥16	≥16	≥16	≥16	4	+		
H459C	26	F	N	Y	≥16	≥16	≥16	0	4	4		+	
H477C	34	F	N	Y	4	≥16	≥16	≥16	≥16	4	+		
H478C	36	F	N	Y	0	≥16	0	≥16	≥16	0	+	+	
H001M	27	M	N	Y	na	na	na	na	na	na		+	
H002M	36	M	N	Y	na	na	na	na	na	na		+	
H003M	29	M	N	Y	na	na	na	na	na	na		+	
H004M	23	M	N	Y	na	na	na	na	na	na		+	
H005M	28	M	N	Y	na	na	na	na	na	na		+	
H006M	33	M	N	Y	na	na	na	na	na	na		+	
H004N	45	M	Y	Y	0	≥16	≥16	0	≥16	4	+	+	
H494C	56	F	Y	Y	na	na	na	na	na	na		+	
H079O	38	F	Y	N	≥16	≥16	0	0	≥16	4	+		
H106S	48	F	Y	N	≥16	≥16	≥16	≥16	≥16	≥16	+	+	
H158O	56	M	Y	N	≥16	≥16	≥16	≥16	≥16	≥16	+		
H170S	40	F	Y	N	≥16	≥16	≥16	≥16	≥16	0	+		
H372C	26	F	Y	N	0	≥16	≥16	≥16	0	4		+	
H455C	25	F	Y	N	≥16	≥16	≥16	0	4	≥16	+	+	
H460C	24	M	Y	N	0	0	≥16	≥16	≥16	0	+		
H463C	29	F	Y	N	≥16	4	≥16	0	≥16	4	+		
H466C	25	F	Y	N	≥16	≥16	≥16	≥16	≥16	≥16	+	+	
H487C	39	M	Y	N	na	na	na	na	na	na	+	+	

Table S2. Summary of T-cell screening and validation experiments. The first columns detail the identity of each peptide. For the screening round, MMr^+CD8^+ T-cell frequencies, MMr^+ counts and percent effector/memory fractions are listed. For the more stringent validation round, donors/peptides with <3 MMr^+ cells counted were excluded from the final analysis, with frequencies noted as 1.0E-7 (i.e. 10^{-7}) and percent effector/memory fractions not assigned (NA). Merged table rows indicate peptides with high sequence homology. Validated epitopes are displayed in red fonts.

HLA	Strain	Position	Sequence	Pipeline	Screening round			Validation round	
					Frequency median	MMr+ counts median (range)	% effector/memory median (range)	Frequency median	% effector/memory median (range)
HLA-A2 n=24 n=5	CVB3	235-244	KLVQRVVYNA	MS	1.6 E-6	1 (0-3)	0 (0-0)		
	CVB3	1038-1048	SILEKSLKALV	MS	1.2 E-6	1 (0-2)	0 (0-100)		
	CVB1/CVB3	1359-1366	KINMPMSV	MS	1.0 E-7	0 (0-1)	0 (0-0)		
	CVB1-2-5	60-68	IMIKSMPAL	In silico	2.2 E-5	30 (8-155)	26 (0-70)	1.0 E-7	NA
	CVB4-6	60-68	VMIKSLPAL	In silico	5.3 E-6	2 (0-5)	20 (20-20)		
	CVB5	169-177	YLGRAGYTV	In silico	5.4 E-5	14 (6-50)	10 (4-55)	7.3 E-6	14 (0-33)
	CVB1-2-4-5-6	271-279	IVMPYINSV	In silico	2.2 E-6	2 (0-24)	85 (0-100)	2.1 E-5	43 (26-82)
	CVB2	291-299	FTLMIIPFV	In silico	1.5 E-6	1 (0-3)	0 (0-100)		
	CVB1-2-3-4-5-6	456-464	AMATGKFLL	In silico	1.0 E-7	0 (0-1)	0 (0-0)		
	CVB1-5	555-563	RMLKDTPFI	In silico	1.0 E-7	0 (0-1)	0 (0-0)		
	CVB2-5	711-719	VLTHQIMYV	In silico	1.0 E-7	0 (0-0)	0 (0-0)		
	CVB4	943-951	YQSHVLLAV	In silico	1.9 E-6	2 (0-5)	0 (0-100)		
	CVB5	945-953	YQTHVLLAV	In silico	5.6 E-5	25 (9-47)	24 (7-55)	6.7 E-6	8 (0-55)
	CVB1-2-3-4-5-6	1246-1254	KLNSSVYSL	In silico	3.0 E-5	13 (1-36)	33 (0-70)	1.3 E-5	38 (0-83)
	CVB1-2-3-4-5-6	1289-1297	QMVSSVDFV	In silico	2.7 E-6	1 (0-5)	40 (40-40)		
	CVB1-2-3-4-6	1307-1315	GILFTSPFV	In silico	1.0 E-7	0 (0-3)	0 (0-33)		
	CVB5	1311-1319	FLFTSPFVL	In silico	8.0 E-6	4 (1-14)	29 (20-56)		
	CVB1-2-3-4-5-6	1509-1517	KLFAGFQGA	In silico	1.0 E-7	0 (0-0)	0 (0-0)		
	CVB1-3	1585-1593	LMNDQEIGV	In silico	7.4 E-7	1 (0-12)	50 (0-100)		
	CVB2-4-5-6	1589-1598	ILMNDQEVGV	In silico	2.5 E-5	11 (1-892)	40 (24-99)	1.0 E-7	NA
	CVB1-5-6	1626-1634	FLAKEEVEV	In silico	1.0 E-7	0 (0-1)	0 (0-0)		
	CVB2-4	1631-1639	FLAREEAEV	In silico	5.3 E-6	2 (0-10)	34 (25-43)		
	CVB1-2-3-4-5-6	1671-1679	RMLMYNFPT	In silico	2.8 E-5	24 (0-154)	25 (0-67)	6.2 E-6	16 (0-33)
	CVB1-2-3-4-5-6	2107-2115	FLVHPVMPM	In silico	1.0 E-7	0 (0-0)	0 (0-0)		
HLA-A3 n=12 n=4	CVB1	804-813	RIYFKPKHVK	MS	4.3 E-5	16 (10-67)	66 (23-88)	1.0 E-7	NA
	CVB1	1132-1140	KIL<u>PEVKEK</u>	MS	2.0 E-5	9 (4-21)	58 (20-71)		
	CVB1 only	*1132-1140*	*KIL <u>PEVKEK</u>*	*MS*				*1.0 E-7*	*24 (20-29)*
	GAD only	*272-280*	*KMF <u>PEVKEK</u>*	*MS*				*1.0 E-7*	*23 (20-55)*
	CVB1 x GAD							1.3 E-5	60 (23-100)
	CVB1/CVB3	1233-1241/1236-1244	SVATNLIGR	MS	5.4 E-6	2 (1-6)	42 (0-83)		
	CVB1/CVB3	1356-1364/1359-1367	KINMPMSVK	MS	1.0 E-5	4 (0-89)	95 (80-100)	3.1 E-5	98 (80-100)
	CVB2	1361-1369	KVNMPMSVK	In silico	1.9 E-5	7 (0-15)	61 (40-100)		
	CVB3	1519-1528	GAYTGVPNQK	MS	1.0 E-7	0 (0-2)	0 (0-0)		
	CVB3	1765-1774	VLRSGDPRLK	MS	1.4 E-5	6 (0-15)	68 (50-75)	1.8 E-5	78 (56-92)
	CVB5	760-768	SMFYDGWAK	In silico	1.9 E-5	12 (2-25)	33 (0-68)	2.9 E-5	33 (9-94)
	CVB1-2-3-4-5-6	1204-1212	KMSNYIQFK	In silico	1.0 E-7	0 (0-0)	0 (0-0)		
	CVB1-2-3-4-5-6	2024-2032	IIIRTLMLK	In silico	4.4 E-6	2 (1-3)	0 (0-0)		
	CVB1-2-3-4-5-6	2027-2035	RTLMLKVYK	In silico	3.0 E-6	1 (0-6)	25 (16-33)		
	CVB2-3-4-5	2161-2169	KIRSVPVGR	In silico	1.0 E-7	0 (0-0)	0 (0-0)		

Table S3. HLA-A2$^+$ CVB-seropositive children recruited for T-cell studies. Columns display study groups, case IDs, donors' demographics, titers of neutralizing antibodies (NAbs) for different CVB serotypes (positive titers of ≥16 in red, borderline titers of 4 in blue; corresponding to serum dilutions of 1/16 and 1/4, respectively), T1D duration and the corresponding symbols used in Fig. S7, with T1D donors contoured in red and healthy donors contoured in blue. Donors recruited in Miami are indicated with black-filled symbols. All other donors were recruited in Paris.

Study group	Case ID	Age (yrs)	Gender (M/F)	CVB1 NAbs	CVB2 NAbs	CVB3 NAbs	CVB4 NAbs	CVB5 NAbs	CVB6 NAbs	T1D duration (yrs)	Symbols
T1D (n=11)	B001M	8	M	≥16	≥16	0	≥16	0	4	3	
	B002M	12	M	0	0	≥16	≥16	0	0	4	
	B003M	12	M	4	≥16	0	0	0	0	7	
	B004M	19	M	0	4	0	≥16	≥16	4	0.6	
	B005M	16	M	4	≥16	≥16	0	≥16	0	0.6	
	B006M	9	F	4	4	0	≥16	≥16	4	6	
	B091S	7	M	na	na	na	na	na	na	0	
	B241R	14	M	≥16	0	≥16	0	0	4	0.02	
	B244V	17	M	0	0	0	≥16	≥16	4	0	
	B252R	11	M	≥16	≥16	≥16	≥16	≥16	≥16	0	
	B234R	11	M	0	≥16	≥16	≥16	≥16	4	0.02	
		12 (7-19)	**91/9%**	**60%**	**70%**	**50%**	**70%**	**60%**	**70%**	**0.6 (0-7)**	
HC (n=10)	HB001M	9	M	na	na	na	na	na	na	na	
	HB002M	18	M	≥16	0	≥16	≥16	0	4	na	
	HB003M	13	F	≥16	≥16	≥16	≥16	4	4	na	
	HB004M	14	M	4	≥16	0	4	4	4	na	
	HB005M	11	M	0	≥16	≥16	≥16	0	0	na	
	HB006M	9	F	4	0	≥16	0	≥16	4	na	
	HB007M	16	M	4	0	≥16	0	≥16	4	na	
	HB030J	8	M	≥16	≥16	≥16	≥16	≥16	≥16	na	
	HB052J	13	M	4	≥16	0	≥16	0	0	na	
	HB055J	13	M	≥16	0	0	≥16	≥16	≥16	na	
		13 (8-18)	**80/20%**	**89%**	**50%**	**60%**	**70%**	**60%**	**70%**	**na**	

Table S4. nPOD spleen, PLN and PBMC specimens analyzed for CVB reactivities. Results available in the nPOD repository are reported for pancreas VP1 staining, Enterovirus qPCR and CVB proteomics (prot) analyses. The last 3 sets of columns detail the number of total $CD8^+$ T cells analyzed and the MMr^+ counts and frequencies for CVB peptides, Flu peptides (Flu MP_{58-66} for HLA-A2, Flu $NP_{265-273}$ for HLA-A3) and naïve viral peptides (HCV $PP_{1406-1415}$ for HLA-A2, HIV nef_{84-92} for HLA-A3). HLA-A2-restricted CVB1-2-3-4-5-$6_{1246-1254}$ and $CVB1_{271-279}$ and HLA-A3-restricted $CVB1_{1356-1364}$/$CVB3_{1359-1367}$ and $CVB1_{1132-1140}$ were analyzed. For $CVB1_{271-279}$ and $CVB1_{1132-1140}$, the corresponding values are marked with an asterisk. Positive results (≥3 MMr^+ counts and $\geq 10^{-5}$ frequency) are shown in bold (and red for CVB). na, not applicable or not available.

	nPOD case RRID SAMN#	Sex	Age (yrs)	T1D (yrs)	Positive aAbs	Pancreas VP1/qPCR/ Prot	HLA	Spleen							PLNs					PBMCs					
								CD8+	CVB MMr+		Flu MMr+		HCV/HIV		CD8+	CVB MMr+		Flu MMr+		CD8+	CVB MMr+		Flu MMr+	
								# ×10^{-3}	#	Freq	#	Freq	#	Freq	# ×10^{-3}	#	Freq	#	Freq	# ×10^{-3}	#	Freq	#	Freq
T1D (n=13)	6052-15879109	M	12	1	IA/IA-2	+/+/−	A2	86.3	1	1.2e-5	0	0	0	0	na	na	na	na	na	na	na	na	na	na
	6211-15879267	F	24	4	A/GAD/IA-2/ZnT	+/+/+	A2 A3	na	na	na	na	na	na	na	38.1	20* 1	5.4e-4* 2.7e-5	8	2.2e-4	46.7	14* 1	3.1e-4* 2.2e-5	14	3.1e-4
	6212-15879268	M	20	5	IA	+/+/−	A2	47.6	0	0	8	1.7e-4	1	2.1e-5	na	na	na	na	na	na	na	na	na	na
	6224-15879280	F	21	2	None	na	A2	30.2	0	0	3	1.0e-4	0	0	na	na	na	na	na	na	na	na	na	na
	6243-15879299	M	13	5	IA	+/na/na	A2	122.8	0	0	43	3.5e-4	2	1.6e-5	na	na	na	na	na	na	na	na	na	na
	6325-15879379	F	20	6	IA/GAD/IA-2	+/+/+	A2	128.4	1	7.8e-6	2	1.6e-5	0	0	na	na	na	na	na	na	na	na	na	na
	6438-15879491	F	39	10	GAD	na	A2	220.2	3	1.4e-5	166	7.5e-4	0	0	na	na	na	na	na	na	na	na	na	na
	6456-15879509	F	30	0	GAD/ZnT8	na	A2	118.2	0	0	22	1.9e-4	0	0	na	na	na	na	na	na	na	na	na	na
	6046-15879103	F	19	8	IA-2/ZnT8	+/+/+	A2 A3	173.0	2 8	1.2e-5 4.6e-5	na 33	na 1.9e-4	na 1	na 5.8e-6	na	na	na	na	na	na	na	na	na	na
	6195-15879251	M	19	5	A/GAD/IA-2/ZnT	−/+/+	A3	73.3	1	1.4e-5	12	1.6e-4	0	0	na	na	na	na	na	na	na	na	na	na
	6265-15879319	M	11	8	IA/GAD	−/−/na	A3	126.2	13	1.0e-4	1	7.9e-6	1	7.9e-6	na	na	na	na	na	na	na	na	na	na
	6362-15879415	M	25	0	GAD	na	A3	40.5	6	1.5e-4	17	4.2e-4	0	0	125.7	29 25*	2.3e-4 2.0e-4*	54	4.3e-4	na	na	na	na	na
	6480-15879533	M	17	3	IA/IA-2	na	A3	17.5	3	1.7e-4	0	0	3	1.7e-4	na	na	na	na	na	na	na	na	na	na
aAb+ (n=10)	6080-15879137	F	69	n/a	IAA/GAD	+/+/na	A2	9.6	0	0	1	1.0e-4	2	2.1e-4	na	na	na	na	na	na	na	na	na	na
	6151-15879207	M	30	n/a	GAD	−/−/−	A2	97.4	1	1.0e-5	10	1.0e-4	0	0	na	na	na	na	na	na	na	na	na	na
	6171-15879227	F	4	n/a	GAD	−/−/na	A2	35.1	4	1.1e-4	0	0	0	0	87.3	1 11*	1.2e-5 1.3e-4*	32	3.7e-4	19.3	9 5*	5.0e-4 2.8e-4*	42	2.2e-3
	6197-15879253	M	22	n/a	GAD/IA-2	+/+/−	A2	87.1	8	9.2e-5	2	2.3e-5	1	1.1e-5	38.1	0 25*	0 6.8e-4	7	1.9e-4	na	na	na	na	na
	6347-15879401	M	9	n/a	IAA	na	A2	97.4	1	1.0e-5	5	5.1e-5	1	1.0e-5	na	na	na	na	na	na	na	na	na	na
	6429-15879482	M	22	n/a	IAA/GAD	−/−/+	A2	41.4	0	0	1	2.4e-5	0	0	na	na	na	na	na	na	na	na	na	na
	6421-15879474	M	7	n/a	GAD	na	A2 A3	129.8	2 2	1.5e-5 1.5e-5	na 2	na 1.5e-5	na 2	na 1.5e-5	na	na	na	na	na	na	na	na	na	na
	6156-15879212	M	40	n/a	GAD	−/na/+	A3	45.3	0	0	13	2.9e-4	1	2.2e-5	na	na	na	na	na	na	na	na	na	na
	6158-15879214	M	40	n/a	IAA/GAD	+/−/+	A3	15.1	0	0	0	0	0	0	na	na	na	na	na	na	na	na	na	na
	6181-15879237	M	32	n/a	GAD	+/na/+	A3	70.0	0	0	61	8.7e-4	1	1.4e-5	na	na	na	na	na	na	na	na	na	na
ND (n=8)	6102-15879159	F	45	n/a	None	+/−/−	A3	19.6	0	0	20	1.0e-3	0	0	na	na	na	na	na	na	na	na	na	na
	6140-15879197	M	38	n/a	None	+/−/−	A3	17.0	0	0	1	5.9e-5	4	2.4e-4	na	na	na	na	na	na	na	na	na	na
	6338-15879392	M	17	n/a	None		A2 A3	134.8	9* 0	6.7e-5* 0	5	3.8e-5	na	na	173.3	3* 2	1.7e-5* 1.2e-5	3	1.7e-5	na	na	na	na	na
	6368-15879421	M	38	n/a	None	na	A3	307.7	10	3.2e-5	972	3.2e-3	2	6.5e-6	na	na	na	na	na	na	na	na	na	na
	6386-15879439	M	14	n/a	None		A2 A3	330.2	48* 4	1.5e-4 1.2e-5	59	1.8e-4	na	na	674.0	54* 9	8.1e-5* 1.4e-5	74	1.1e-4	na	na	na	na	na
	6420-15879473	M	11	n/a	None		A2 A3	394.3	18*1	4.6e-5* 2.6e-6	19	4.9e-5	na	na	na	na	na	na	na	120.0	9* 0	7.6e-5* 0	7	5.9e-5
	6461-15879514	M	14	n/a	None		A2 A3	75.0	44*3	6.0e-4* 4.1e-5	9	1.2e-4	na	na	37.3	33* 2	9.1e-4* 5.5e-5	11	3.0e-4	na	na	na	na	na
	6488-15879541	F	5	n/a	None		A2 A3	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

Table S5. MMr panels used for the analysis of nPOD spleen, PLN and PBMC specimens. The combined HLA-A2/HLA-A3 MMr panels were used for specimens from HLA-A2/A3^{+} nPOD donors. The homologous $CVB1_{1132\text{-}1140}$/ $GAD_{272\text{-}280}$ peptides are marked with an asterisk. The MMr panel for PLN and PBMC specimens was also used for splenocytes from nPOD donors 6338, 6386, 6420, 6461 and 6480. The extended phenotypic panel used for PLN and PBMC specimens is also detailed.

Spleen specimens (Fig. 4A-B)		
MMr panel	**Fluorochrome pair**	**Peptide**
HLA-A2	PE / APC	HCV PP 1406-1415
	PE / BV711	CVB5 169-177
	PE / BV786	CVB5 945-953
	APC / BV711	CVB1-2-3-4-5-6 1246-1254
	APC / BV786	CVB1-2-4-5-6 271-279
	BV711 / BV786	Flu MP 58-66
HLA-A3	PE / APC	HIV nef 84-92
	PE / BV711	GAD 272-280*
	PE / BV786	CVB1 1356-1364 / CVB3 1359-1367
	APC / BV711	CVB1 1132-1140*
	APC / BV786	CVB5 760-768
	BV711 / BV786	Flu NP 265-273
HLA-A2 + HLA-A3	PE / APC	HLA-A3 HIV nef 84-92
	PE / BV711	HLA-A3 GAD 272-280
	PE / BV786	HLA-A3 CVB1 1356-1364 / CVB3 1359-1367
	APC / BV711	HLA-A2 CVB1-2-3-4-5-6 1246-1254
	APC / BV786	HLA-A2 CVB1-2-4-5-6 271-279
	BV711 / BV786	HLA-A3 Flu NP 265-273
PLN and PBMC specimens (Fig. 4C-D-E)		
MMr panel	**Fluorochrome pair**	**Peptide**
HLA-A2	PE/BV650	CVB1-2-3-4-5-6 1246-1254
	APC/PE-CF594	CVB1 271-279
	BV711	Flu MP 58-66
HLA-A3	PE/BV650	CVB1 1356-1364 / CVB3 1359-1367
	APC/PE-CF594	CVB1 1132-1140
	BV711	Flu NP 265-273
HLA-A2 + HLA-A3	PE/BV650	HLA-A2 CVB1 271-279
	APC/PE-CF594	HLA-A3 CVB1 1356-1364 / CVB3 1359-1367
	BV711	HLA-A3 Flu NP 265-273
Phenotype markers	Far Red	Live/Dead
	APC-H7	CD3
	PE-Cy7	CD8
	FITC	CD45RA
	BV421	CCR7
	BV510	CD25
	BV605	PD-1

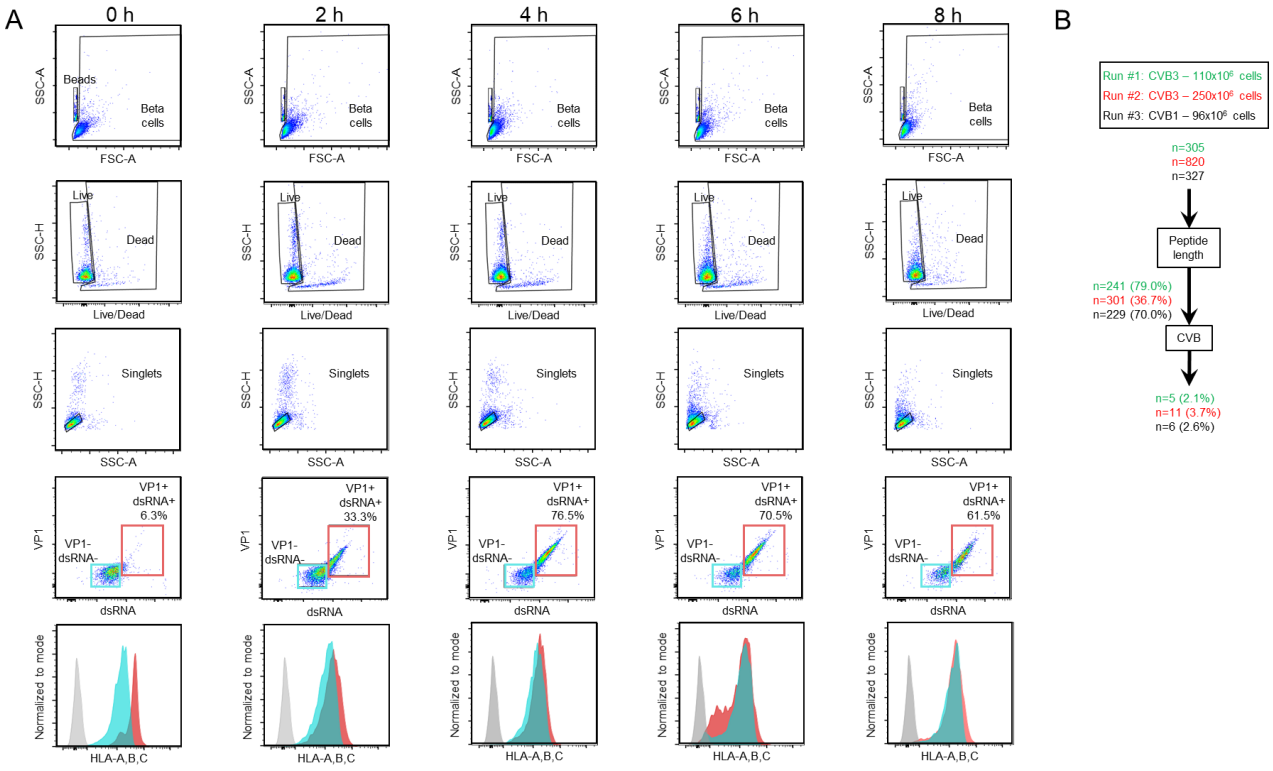


Figure S1. Kinetics of CVB3 infection in ECN90 β cells. A. Representative flow cytometry dot plots and gating strategy. After gating on physical parameters and viable (Live/Dead$^-$) singlets VP1$^+$dsRNA$^+$ cells were visualized, HLA-I and PD-L1 staining in CVB3-infected ECN90 β cells (300 MOI). After gating on physical parameters and on viable (Live/Dead$^-$) cells, VP1$^+$dsRNA$^+$ were visualized, and their HLA-I expression compared with that of the VP1$^-$dsRNA$^-$ population, as shown in the last row (with isotype control-stained cells overlaid in grey). A fixed amount of beads was added to each well before staining to normalize cell counts. **B.** The 6-h time point was chosen and 3 HLA-I peptidomics experiments performed (color code). Cells were lysed and HLA-I-bound peptides eluted and sequenced by MS. Sequences were filtered for peptide length (8-12 aa) and for matches with the translated RNA sequences of the CVB strains used. The number of peptides retained at each filtering step are shown (percent of peptides out of those from the previous step in parentheses).



Capsid protein P1 (VP1-VP4)

Protein P2 (2A-C)

Protein P3 (3A-D)

ABCD HLA-eluted peptides (CVB1, CVB3)

D aa identities in other serotypes

ABCD *In silico* predicted peptides

D aa identities in other serotypes

Overlapping predictions

In vitro binding assay passed

T-cell screening passed

T-cell screening and validation passed

X Residue difference between RNA-sequenced strains and reference serotypes

Figure S2. Mapping of HLA-eluted and *in silico* predicted peptides and conservation across CVB serotypes. ECN90 β cells were infected with either CVB1 or CVB3. The pHLA-I complexes were purified and peptides sequenced by MS (bold fonts, highlighted in yellow). A parallel *in silico* search across relevant strains (see Methods) identified nonamer peptides predicted to bind HLA-A2 or HLA-A3 (bold font, highlighted in grey). Within these peptides, the aa conserved across serotypes are highlighted in yellow (for HLA-eluted peptides) or grey (for *in silico* predicted peptides). All the peptides were tested for their ability to bind to the corresponding HLA-I monomers (Fig. S3). Binding peptides thus identified (encircled with blue dashes) were screened by combinatorial HLA-I multimer (MMr) assays. Those scored as MMr^{+} in this screening phase are encircled with red dashes. Peptides whose T-cell recognition was confirmed in validation experiments are encircled with continuous red lines, with their aa positions indicated on top. CVB1 and CVB3 aa sequences are those translated from the RNA-sequencing of strains used for *in vitro* infection. The aa differences between the experimental (RNA-sequenced) strains used for HLA-I peptidomics and the reference sequences used for *in-silico* predictions are indicated in red letters.

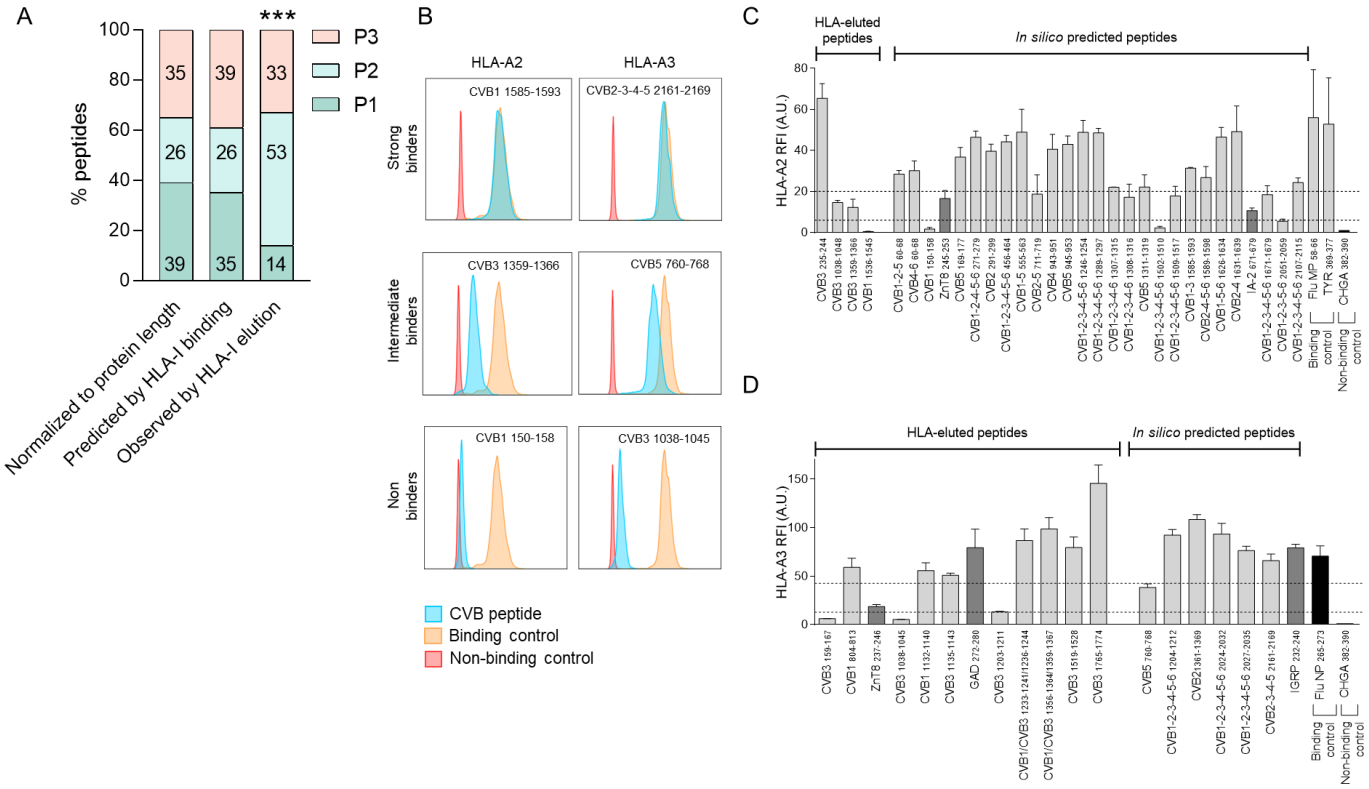


Figure S3. Observed vs. predicted HLA-I-binding peptides within CVB proteins and HLA-I binding measurements. A. Distribution of HLA-I-eluted vs. predicted HLA-I-binding peptides across the P1, P2 and P3 CVB proteins. The graph shows the percent distribution of peptides expected within P1, P2 and P3 based on the aa length (first bar) and on the number of unique predicted binders of each protein (second bar; using NetMHCcons 1.1) compared to the number of unique observed binders experimentally eluted (third bar). Reference CVB1/CVB3 aa sequences were those translated from the RNA-sequencing of strains used for *in vitro* infection. ***p<0.0001 by χ^2 test. **B-D.** HLA-I binding measurements. Biotinylated recombinant HLA-A2 and HLA-A3 molecules were folded with the indicated peptides and β2-microglobulin. Complexes were captured on streptavidin-coated beads and stained for β2-microglobulin. **B.** Representative staining for HLA-A2 (left) and HLA-A3 complexes (right) folded with CVB peptides (blue profiles) displaying strong (top), intermediate (middle) and no binding (bottom). Binding controls (Flu MP_{58-66} and $TYR_{369-377}$ for HLA-A2; Flu $NP_{265-273}$ for HLA-A3; yellow profiles) and non-binding controls ($CHGA_{382-390}$; red profiles) are shown for comparison. **C-D.** HLA-eluted and *in-silico* predicted peptides were tested for their binding to HLA-A2 (C) and HLA-A3 (D). Horizontal dotted lines indicate cut-off values for scoring strong, intermediate and non-binders. Bars represent mean ± SEM values from 2-3 different experiments of the relative fluorescence intensity (RFI), i.e. the median fluorescence intensity

arbitrary units (A.U.) normalized to the $CHGA_{382-390}$ non-binding peptide. *In-silico* predicted HLA-A2-binding peptides $CVB1_{150-158}$ (KLPDALSQM; and its binding homologous $ZnT8_{245-253}$, VMGDALGSV), CVB1-2-3-4-5-$6_{1502-1510}$ (GIIYIIYKL; IEDB #20311) and CVB1-2-3-5-$6_{2051-2059}$ (VIASYPWPI) did not confirm as binders and are therefore not listed in Fig. 1F.

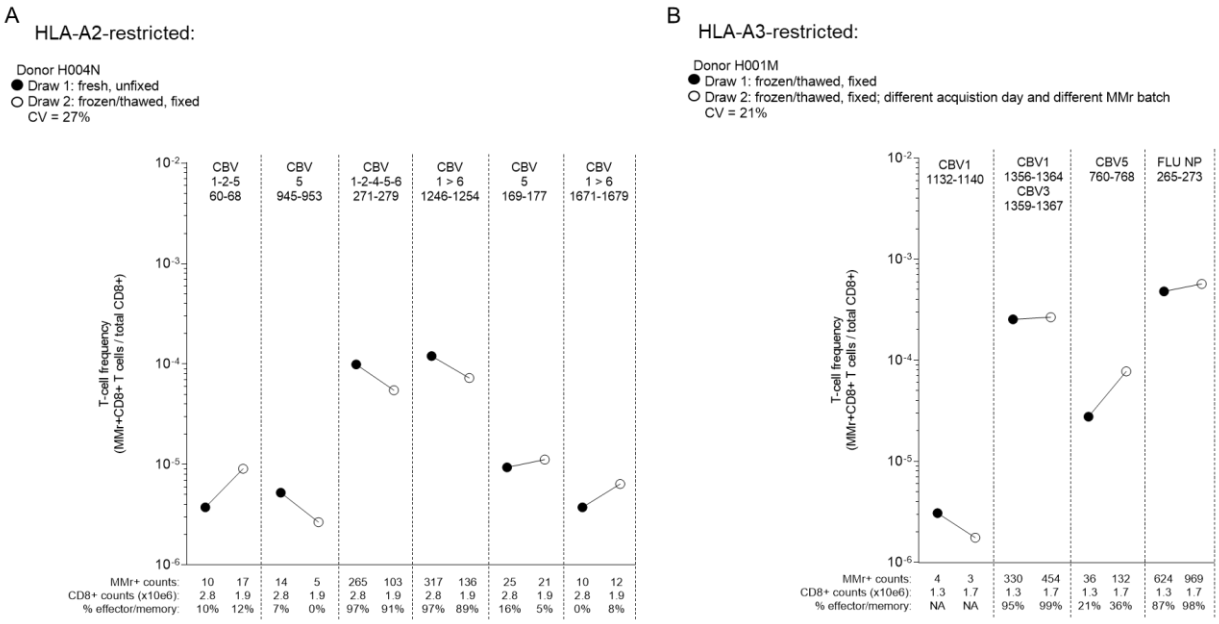


Figure S4. Reproducibility of HLA-A2 and HLA-A3 MMr assays using CVB peptides. A. Reproducibility using CVB peptide-loaded HLA-A2 MMrs. Donor H004N was analyzed using two separate blood draws and the same MMr batch, with PBMCs stained freshly after isolation and not fixed (black dots) or stained on frozen/thawed samples fixed after staining (white dots). Coefficient of variation (CV) of measured MMr^+CD8^+ T-cell frequencies was 27%. **B.** Reproducibility using CVB peptide-loaded HLA-A3 MMrs. Donor H001M was analyzed using frozen/thawed PBMCs from two separate blood draws and two separate MMr batches, with PBMCs fixed after staining. CV=21%. MMr^+ counts, the number of total $CD8^+$ T cells acquired and the percent effector/memory fraction among MMr^+ cells are shown below each symbol. NA, not applicable (effector/memory fraction not assigned when <5 MMr^+ cells detected).

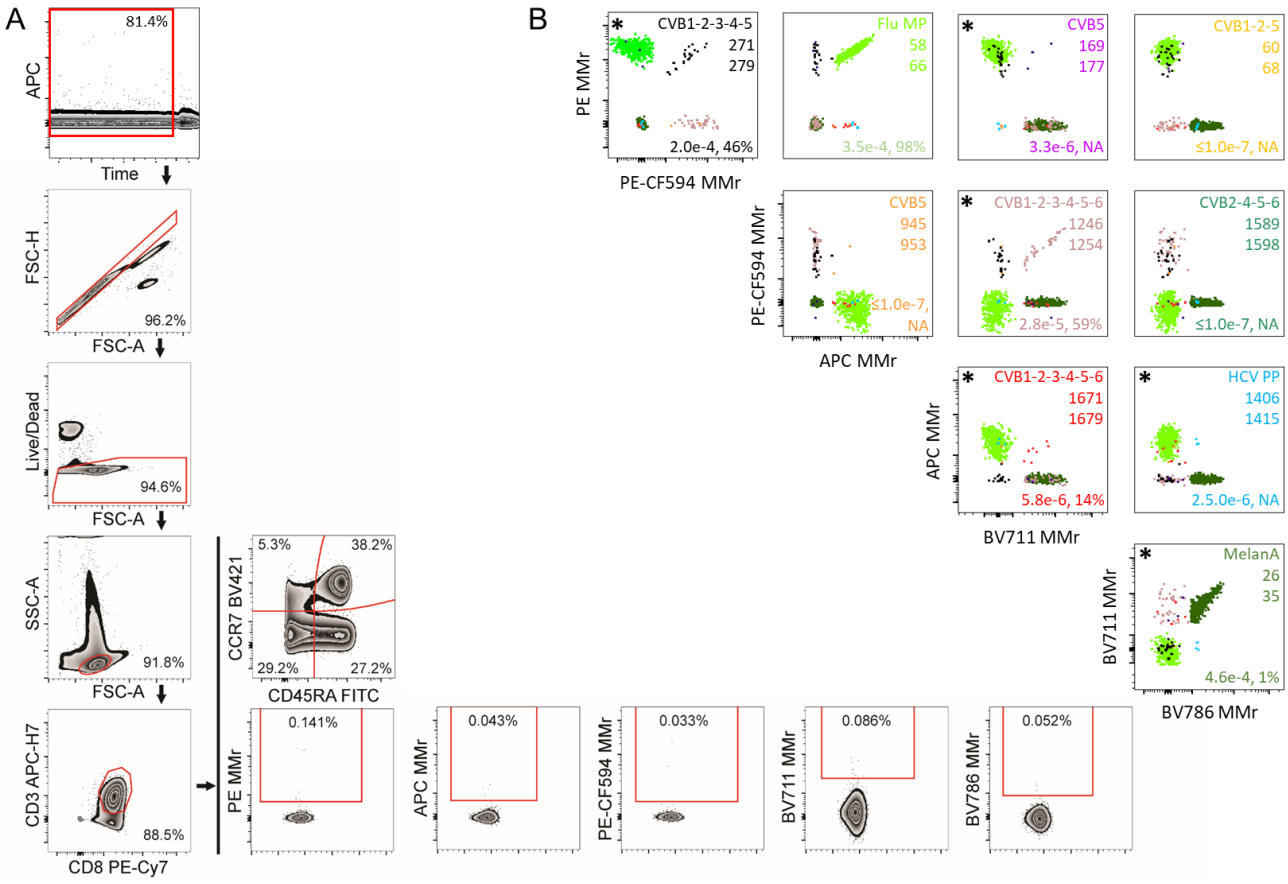


Figure S5. Representative staining with the HLA-A2 MMr panel used for T-cell validation experiments. Frozen-thawed PBMCs from donor H004N (HLA-A2/A3^{+}) were magnetically depleted of CD8^{-} cells before staining, acquisition, and analysis. **A.** Gating strategy. After verifying the stability of each fluorescence signal over the acquisition time (example shown for APC), cells were sequentially gated on singlets (FSC-A/FSC-H dot plot), live cells (Live/Dead Aqua^{-}), small lymphocytes (FSC-A/SSC-A dot plot), CD3^{+}CD8^{+} T cells, and total PE^{+}, APC^{+}, PE-CF594^{+}, BV711^{+}, and BV786^{+} MMr^{+} cells. Using Boolean operators, these latter gates allowed to selective visualize each double-MMr^{+} population by including only those events positive for the corresponding fluorochrome pair. The CD45RA/CCR7 staining distribution of total CD3^{+}CD8^{+} T cells is also shown. **B.** The final readout obtained is shown for the 10 peptides analyzed. Events corresponding to each epitope-reactive population are overlaid in different colors within each dot plot. Peptides scoring positive for this donor are marked with an asterisk. Numbers in each panel indicate the MMr^{+}CD8^{+} T-cell frequency out of total CD8^{+} T cells and the percent effector/memory fraction among MMr^{+} cells (NA when not assigned, i.e. <5 MMr^{+} cells counted).

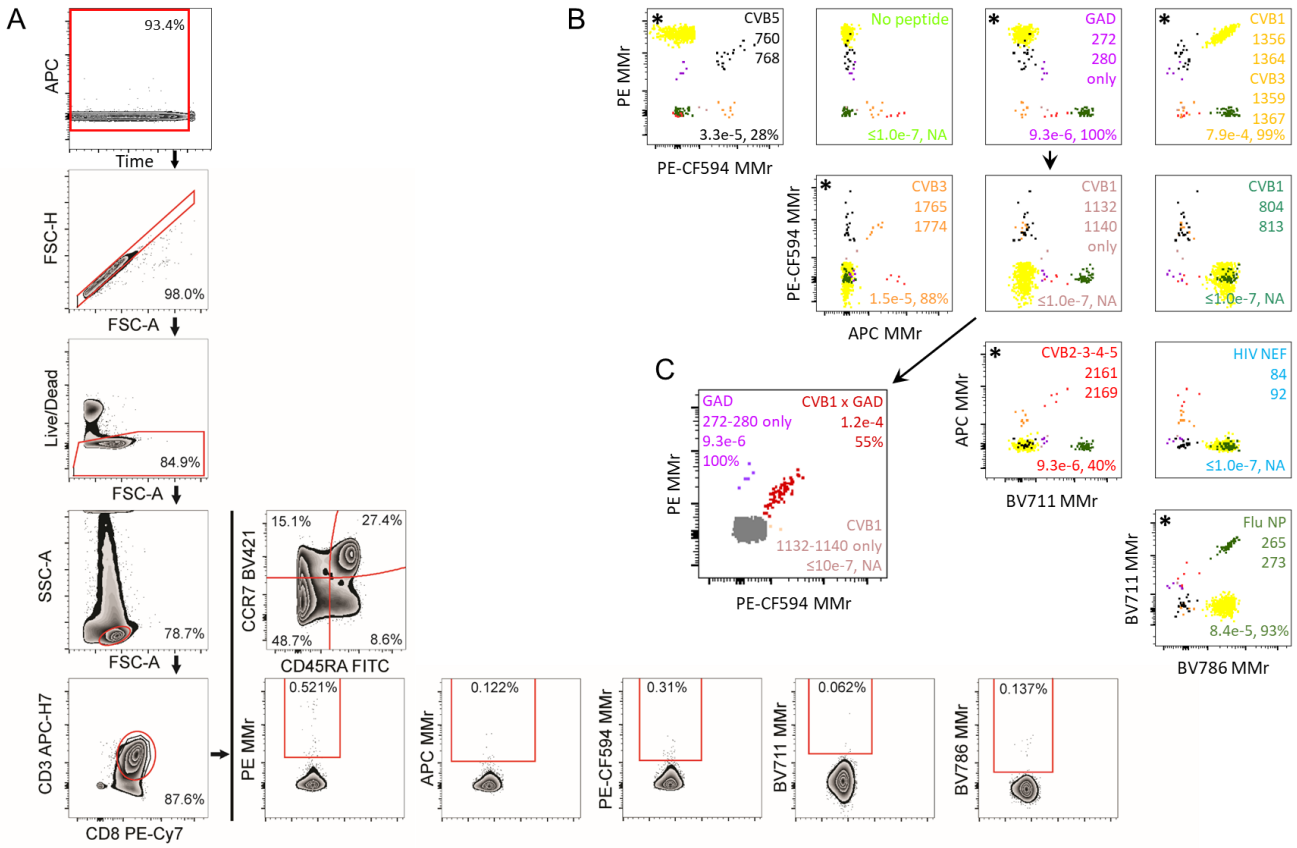


Figure S6. Representative staining with the HLA-A3 MMr panel used for T-cell validation experiments. Frozen-thawed PBMCs from donor H004N (HLA-A2/A3^{+}) were magnetically depleted of CD8^{-} cells before staining, acquisition, and analysis. **A.** Gating strategy. After verifying the stability of each fluorescence signal over the acquisition time (example shown for APC), cells were sequentially gated and Boolean operators applied as in Fig. S5. **B.** The final readout obtained is shown for the 9 peptides analyzed (with the PE/APC fluorochrome pair left empty). Data representation is the same as in Fig. S5. **C.** The gating strategy based on positivity for 2 MMr fluorochromes precludes visualization of CVB1$_{1132\text{-}1140}$/GAD$_{272\text{-}280}$ cross-reactive CD8^{+} T cells, since they are positive for 3 fluorochromes (PE, PE-CF594 and the shared BV711 fluorochrome). Hence, one further gating was performed for these fractions by selecting cells positive for only 2 fluorochromes (PE-CF594/BV711^{+}, PE^{-}, labeled as "CVB1$_{1132\text{-}1140}$ only"; or PE/BV711^{+}, PE-CF594^{-}, labeled as "GAD$_{272\text{-}280}$ only") or for all 3 (PE/PE-CF594/BV711^{+}, labeled as "CVB1 x GAD"). The resulting overlaid dot plot is displayed (with cells negative for all 3 fluorochromes depicted in grey), showing that most T cells recognizing these CVB1$_{1132\text{-}1140}$ and/or GAD$_{272\text{-}280}$ epitopes are cross-reactive (shown in red).

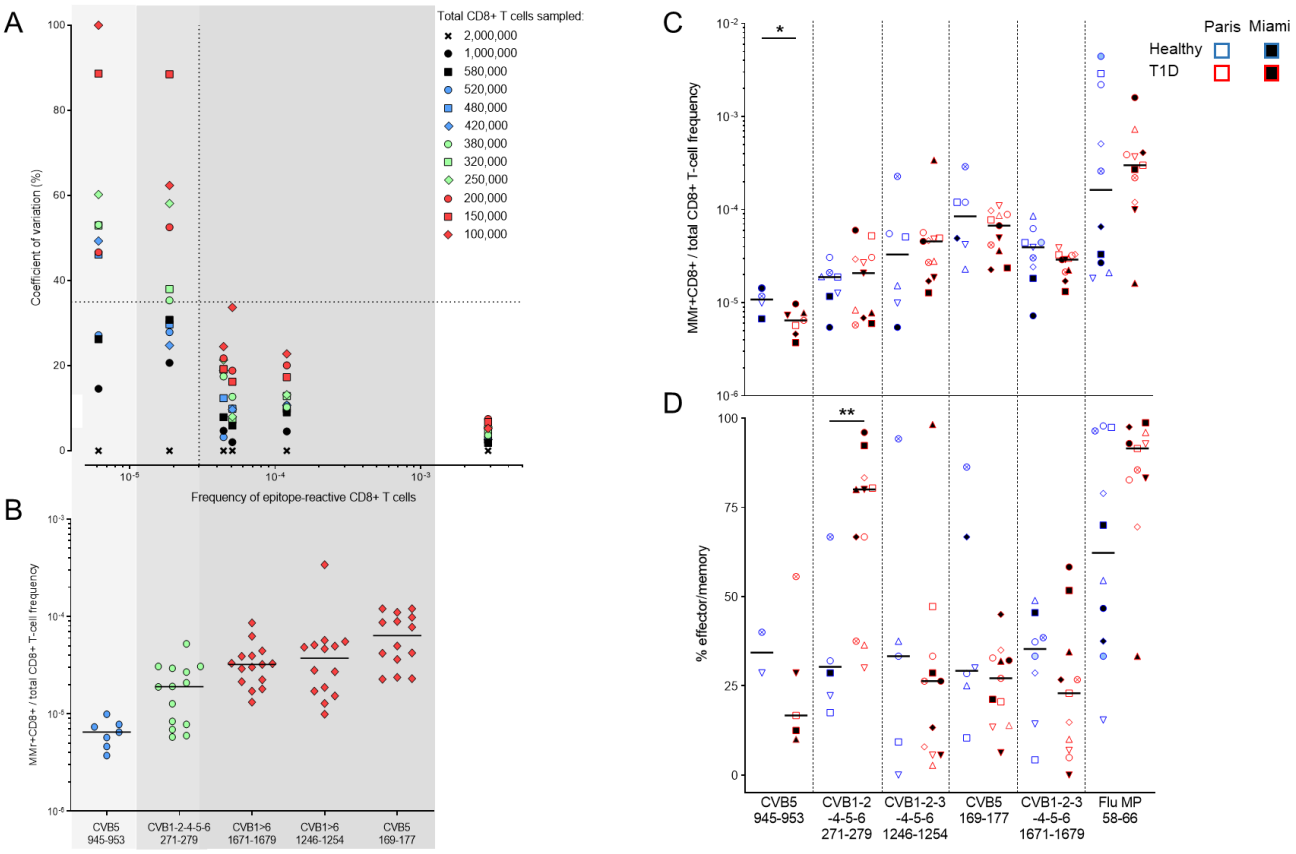


Figure S7. Circulating HLA-A2-restricted CVB-reactive CD8$^+$ T cells in pediatric donors. **A-B.** Definition of the minimal CD8$^+$ T-cell sampling size required to accurately measure CVB epitope-reactive CD8$^+$ T-cell frequencies. In panel A, decreasing numbers of total CD8$^+$ T cells from a single blood draw were acquired for donor H004N after staining with CVB peptide-loaded HLA-A2 MMrs. The MMr$^+$CD8$^+$ T-cell frequency measured on the maximal number (2×10^6) of total CD8$^+$ T cells acquired was considered as the gold standard (indicated with crossed symbols), and the CV of frequencies measured with lower numbers of acquired T cells calculated accordingly. For CVB epitopes recognized by MMr$^+$CD8$^+$ T-cells showing a mean frequency >$3/10^5$ (vertical dotted line), the acquisition of a minimal number of 100,000 total CD8$^+$ T cells (red diamonds) yields an accurate frequency measurement (CV<35%; horizontal dotted line). For MMr$^+$CD8$^+$ T-cell frequencies <$2/10^5$, the acquisition of a minimal number of 380,000 total CD8$^+$ T cells (green circles) is required to maintain a CV<35%. For MMr$^+$CD8$^+$ T-cell frequencies <$7/10^6$, the acquisition of a minimal number of 520,000 total CD8$^+$ T cells (blue circles) is required to maintain a CV<35%. Panel B depicts the median frequency and distribution of MMr$^+$CD8$^+$ T cells using the indicated CVB peptides in pediatric donors with >500,000 total CD8$^+$ T cells acquired. Based on these frequencies, the cut-offs defined in panel A for the minimal number of total CD8$^+$ T cells acquired were applied for each peptide, as

indicated by grey shades: ≥520,000 (blue circles) for $CVB5_{945-953}$, ≥380,000 (green circles) for $CVB1\text{-}2\text{-}4\text{-}5\text{-}6_{271-279}$, ≥100,000 (red diamonds) for all other epitopes. Using these cut-offs for each peptide of interest, only those samples allowing for an accurate frequency measurement were retained for final analysis. **C-D.** Frequency (C) and percent effector/memory phenotype (D) of $MMr^{+}CD8^{+}$ T cells recognizing the indicated CVB epitopes or a control memory Flu MP_{58-66} epitope in T1D (blue) and healthy (red) pediatric donors (listed in Table S3). Data representation is the same as in Fig. 3. Children recruited in Paris and Miami are represented with symbols filled in white and black, respectively. *p=0.04 and **p=0.01 by Mann-Whitney U test. The HLA-A2 MMr staining panel was as follows: $CVB1\text{-}2\text{-}3\text{-}4\text{-}5\text{-}6_{1671-1679}$ PE/APC, $CVB5\text{-}6_{169-177}$ PE/BV711, $CVB5_{945-953}$ PE/BV786, $CVB1\text{-}2\text{-}3\text{-}4\text{-}5\text{-}6_{1246-1254}$ APC/BV711, $CVB1\text{-}2\text{-}4\text{-}5\text{-}6_{271-279}$ APC/BV786, Flu MP_{58-66} BV711/BV786.

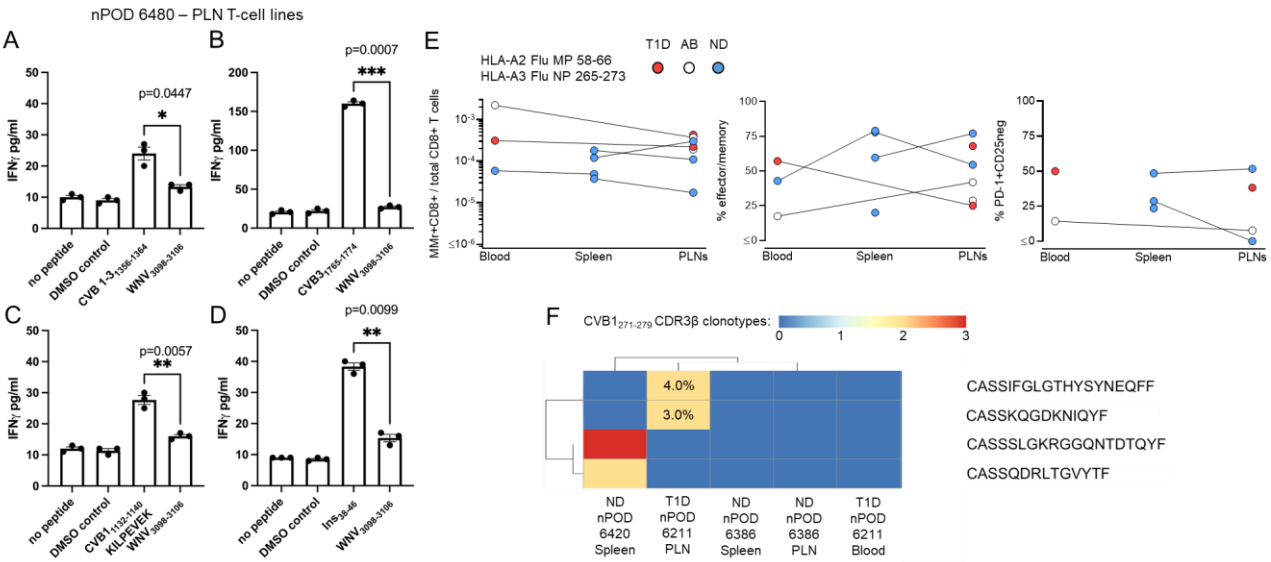


Figure S8. PLN T cells in HLA-A3$^+$ T1D nPOD case 6480 and expanded CVB1$_{271-279}$ CDR3β clonotypes in lymphoid tissues from nPOD donors. A-D. PLN CD8$^+$ T cells were stimulated with the indicated CVB peptides for 14 days, followed by a 48-h recall with the stimulating peptide, the negative control WNV$_{3098-3106}$ peptide or DMSO diluent and cytokine measurement on culture supernatants. Besides IFN-γ, TNF-α and IL-6 were also secreted in response to CVB1$_{1765-1774}$ peptide; no reactivity to CVB1$_{804-813}$ was detected (not shown). **E.** Influenzavirus epitope-reactive CD8$^+$ T cells in lymphoid tissues from nPOD cases (see details in Table S4). MMr$^+$CD8$^+$ T-cell frequencies, percent effector/memory MMr$^+$ and PD-1$^+$CD25$^-$ effector/memory MMr$^+$ cells (for donors/epitopes with cell counts ≥5) are depicted across available tissues in the left, middle and right panel, respectively. **F.** Expanded TCR CDR3β clonotypes in individual CBV1$_{271-279}$-reactive CD8$^+$ T cells from nPOD cases/tissues depicted in Fig. 4C-D-E-F. Expanded clonotypes were defined as those found in at least 2 MMr$^+$ cells in the same tissue from the same donor, and are clustered according to frequency (one columns per tissue/case) and sequence similarity (one row per CDR3β clonotype). For the PLN CD8$^+$ T cells of T1D nPOD case 6211 (second column), percentages indicate the CDR3β clonotype frequency out of total PLN CD8$^+$ T cells sequenced by Seay et al., available at http://clonesearch.jdrfnpod.org/.

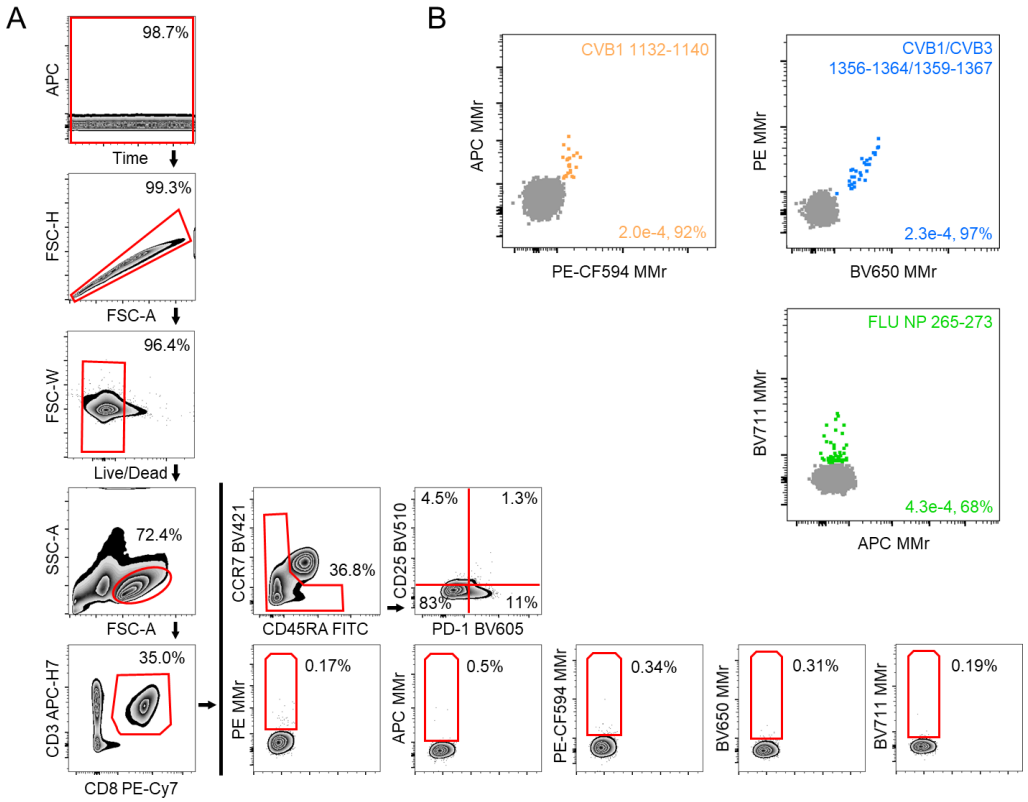


Figure S9. Representative HLA-A3 MMr staining on PLN cells from T1D nPOD case 6362. Given the limited cell numbers available, frozen-thawed PLN cells were directly stained and acquired, without prior magnetic depletion of $CD8^-$ cells. **A.** Gating strategy. After verifying the stability of each fluorescence signal over the acquisition time (example shown for APC), cells were sequentially gated on singlets (FSC-A/FSC-H dot plot), live cells (Live/Dead $Aqua^-$), small lymphocytes (FSC-A/SSC-A dot plot), $CD3^+CD8^+$ T cells, and total PE^+, APC^+, $PE\text{-}CF594^+$, $BV650^+$ and $BV711^+$ MMr^+ cells. Using Boolean operators, these latter gates allowed to selective visualize each double-MMr^+ population by including only those events positive for the corresponding fluorochrome pair. Contrary to previous staining panels, each fluorochrome was here used for only one MMr, in order to allow sorting of individual CVB double-MMr^+ cells for TCR sequencing, with control Flu MMrs labeled with a single BV711 fluorochrome. The CD45RA/CCR7 staining on total $CD3^+CD8^+$ T cells and CD25/PD-1 staining on the effector/memory gate displayed is also shown. **B.** The final readout obtained is shown for the 3 peptides analyzed. Each dot plot displays a color-coded overlay of individual double-MMr^+ subsets to visualize the separation of each epitope-reactive $CD8^+$ T-cell fraction relative to the MMr^- population (gray). Numbers in each panel indicate the MMr^+CD8^+ T-cell frequency out of total $CD8^+$ T cells and the percent effector/memory fraction among MMr^+ cells. PD-1^+ cells (not shown) were gated on MMr^+ effector/memory fractions.

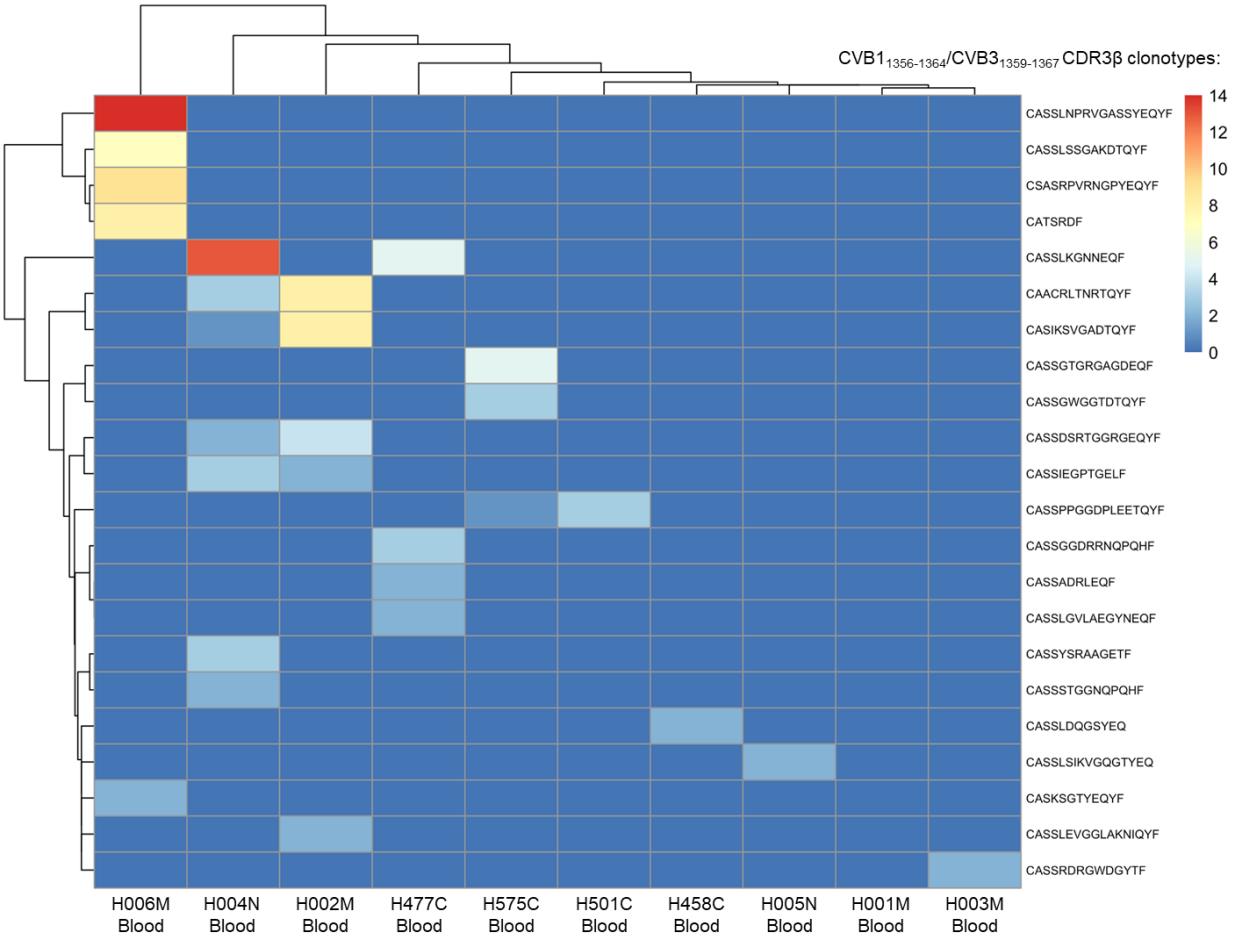


Figure S10. Expanded TCR CDR3β clonotypes in individual CBV1/CVB3$_{1356-1364/1359-1367}$-reactive CD8$^+$ T cells from the blood of living CVB-seropositive healthy donors. Expanded clonotypes were defined as those found in at least 2 MMr$^+$ cells in the same donor, and are clustered according to frequency (one columns per donor) and sequence similarity (one row per CDR3β clonotype).

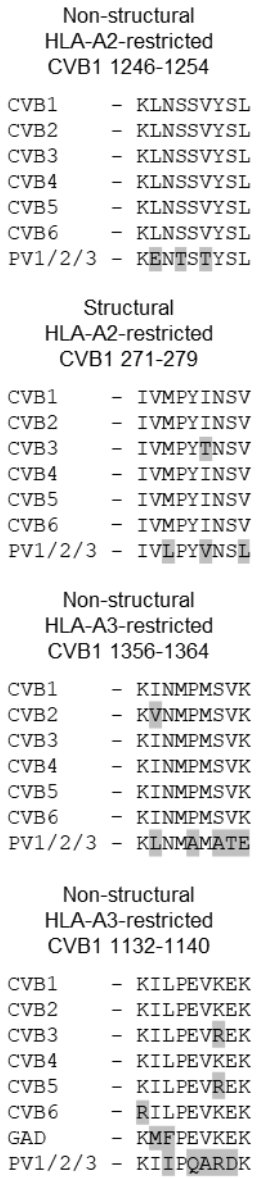


Figure S11. Alignment of the main CD8$^+$ T-cell epitopes identified across CVB and Poliovirus serotypes. Amino-acid differences are shaded in grey. For Poliovirus, the 3 serotypes included in the inactivated poliovirus vaccine were used as reference sequences: serotype 1 (Mahoney), 2 (MEF1) and 3 (Saukett). The alignment refers to the serotype with the closest match.

References

1. Gillespie KM, Bain SC, Barnett AH, Bingley PJ, Christie MR, Gill GV, Gale EA: The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. Lancet 2004;364:1699-1700

2. Patterson CC, Harjutsalo V, Rosenbauer J, Neu A, Cinek O, Skrivarhaug T, Rami-Merhar B, Soltesz G, Svensson J, Parslow RC, Castell C, Schoenle EJ, Bingley PJ, Dahlquist G, Jarosz-Chobot PK, Marciulionyte D, Roche EF, Rothe U, Bratina N, Ionescu-Tirgoviste C, Weets I, Kocova M, Cherubini V, Rojnic Putarek N, deBeaufort CE, Samardzic M, Green A: Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989-2013: a multicentre prospective registration study. Diabetologia 2019;62:408-417

3. Ziegler AG, Bonifacio E, Group B-BS: Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes. Diabetologia 2012;55:1937-1943

4. Parikka V, Nanto-Salonen K, Saarinen M, Simell T, Ilonen J, Hyoty H, Veijola R, Knip M, Simell O: Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk. Diabetologia 2012;55:1926-1936

5. Carré A, Vecchio F, Flodstrom-Tullberg M, You S, R. M: From association to causality: possible diabetogenic mechanisms of Coxsackievirus B infection and implications for type 1 diabetes prevention. Endocrine Rev, under review.

6. Ifie E, Russell MA, Dhayal S, Leete P, Sebastiani G, Nigi L, Dotta F, Marjomaki V, Eizirik DL, Morgan NG, Richardson SJ: Unexpected subcellular distribution of a specific isoform of the Coxsackie and adenovirus receptor, CAR-SIV, in human pancreatic beta cells. Diabetologia 2018;61:2344-2355

7. Laitinen OH, Honkanen H, Pakkanen O, Oikarinen S, Hankaniemi MM, Huhtala H, Ruokoranta T, Lecouturier V, Andre P, Harju R, Virtanen SM, Lehtonen J, Almond JW, Simell T, Simell O, Ilonen J, Veijola R, Knip M, Hyoty H: Coxsackievirus B1 Is Associated With Induction of beta-Cell Autoimmunity That Portends Type 1 Diabetes. Diabetes 2014;63:446-455

8. Vehik K, Lynch KF, Wong MC, Tian X, Ross MC, Gibbs RA, Ajami NJ, Petrosino JF, Rewers M, Toppari J, Ziegler AG, She JX, Lernmark A, Akolkar B, Hagopian WA, Schatz DA, Krischer JP, Hyoty H, Lloyd RE, Group TS: Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. Nat Med 2019;25:1865-1872

9. Dotta F, Censini S, van Halteren AG, Marselli L, Masini M, Dionisi S, Mosca F, Boggi U, Muda AO, Prato SD, Elliott JF, Covacci A, Rappuoli R, Roep BO, Marchetti P: Coxsackie B4 virus infection of beta cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. Proc Natl Acad Sci USA 2007;104:5115-5120

10. Krogvold L, Edwin B, Buanes T, Frisk G, Skog O, Anagandula M, Korsgren O, Undlien D, Eike MC, Richardson SJ, Leete P, Morgan NG, Oikarinen S, Oikarinen M, Laiho JE, Hyoty H, Ludvigsson J, Hanssen KF, Dahl-Jorgensen K: Detection of a low-grade enteroviral infection

in the islets of langerhans of living patients newly diagnosed with type 1 diabetes. Diabetes 2015;64:1682-1687

11. Dunne JL, Richardson SJ, Atkinson MA, Craig ME, Dahl-Jorgensen K, Flodstrom-Tullberg M, Hyoty H, Insel RA, Lernmark A, Lloyd RE, Morgan NG, Pugliese A: Rationale for enteroviral vaccination and antiviral therapies in human type 1 diabetes. Diabetologia 2019;62:744-753

12. Stone VM, Hankaniemi MM, Laitinen OH, Sioofy-Khojine AB, Lin A, Diaz Lozano IM, Mazur MA, Marjomaki V, Lore K, Hyoty H, Hytonen VP, Flodstrom-Tullberg M: A hexavalent Coxsackievirus B vaccine is highly immunogenic and has a strong protective capacity in mice and nonhuman primates. Sci Adv 2020;6:eaaz2433

13. Ashton MP, Eugster A, Walther D, Daehling N, Riethausen S, Kuehn D, Klingel K, Beyerlein A, Zillmer S, Ziegler AG, Bonifacio E: Incomplete immune response to coxsackie B viruses associates with early autoimmunity against insulin. Sci Rep 2016;6:32899

14. Croft NP, Smith SA, Wong YC, Tan CT, Dudek NL, Flesch IEA, Lin LCW, Tscharke DC, Purcell AW: Kinetics of Antigen Expression and Epitope Presentation during Virus Infection. PLOS Pathogens 2013;9:e1003129

15. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, Lalanne AI, Sebastiani G, Carre A, Pinto S, Culina S, Corcos N, Bugliani M, Marchetti P, Armanet M, Diedisheim M, Kyewski B, Steinmetz LM, Buus S, You S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Dotta F, Scharfmann R, Eizirik DL, Verdier Y, Vinh J, Mallone R: Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 2018;28:946-960 e946

16. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, Kuranda K, Nigi L, Eugster A, Osterbye T, Maugein A, McLaren JE, Ladell K, Larger E, Beressi JP, Lissina A, Appay V, Davidson HW, Buus S, Price DA, Kuhn M, Bonifacio E, Battaglia M, Caillat-Zucman S, Dotta F, Scharfmann R, Kyewski B, Mallone R, ImMaDiab Study G: Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013

17. Azoury ME, Tarayrah M, Afonso G, Pais A, Colli ML, Maillard C, Lavaud C, Alexandre-Heymann L, Gonzalez-Duque S, Verdier Y, Vinh J, Pinto S, Buus S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Eizirik DL, You S, Mallone R: Peptides Derived From Insulin Granule Proteins are Targeted by CD8+ T Cells Across MHC Class I Restrictions in Humans and NOD Mice. Diabetes 2020;69:2678-2690

18. Azoury ME, Samassa F, Buitinga M, Nigi L, Brusco N, Callebaut A, Giraud M, Irla M, Lalanne AI, Carre A, Afonso G, Zhou Z, Brandao B, Colli ML, Sebastiani G, Dotta F, Nakayama M, Eizirik DL, You S, Pinto S, Mamula MJ, Verdier Y, Vinh J, Buus S, Mathieu C, Overbergh L, Mallone R: CD8+ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes. Diabetes 2021;70:2879-2891

19. Alanio C, Lemaitre F, Law HK, Hasan M, Albert ML: Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. Blood 2010;115:3718-3725

20. Anderson AM, Landry LG, Alkanani AA, Pyle L, Powers AC, Atkinson MA, Mathews CE, Roep BO, Michels AW, Nakayama M: Human islet T cells are highly reactive to preproinsulin in type 1 diabetes. Proc Natl Acad Sci U S A 2021;118:e2107208118

21. Seay HR, Yusko E, Rothweiler SJ, Zhang L, Posgai AL, Campbell-Thompson M, Vignali M, Emerson RO, Kaddis JS, Ko D, Nakayama M, Smith MJ, Cambier JC, Pugliese A, Atkinson MA, Robins HS, Brusko TM: Tissue distribution and clonal diversity of the T and B cell repertoire in type 1 diabetes. JCI insight 2016;1:e88242

22. Oikarinen M, Laiho JE, Oikarinen S, Richardson SJ, Kusmartseva I, Campbell-Thompson M, Morgan NG, Pugliese A, Tauriainen S, Toniolo A, Hyöty H: Detection of enterovirus protein and RNA in multiple tissues from nPOD organ donors with type 1 diabetes. bioRxiv 2018:459347

23. Mann SE, Zhou Z, Landry LG, Anderson AM, Alkanani AK, Fischer J, Peakman M, Mallone R, Campbell K, Michels AW, Nakayama M: Multiplex T Cell Stimulation Assay Utilizing a T Cell Activation Reporter-Based Detection System. Front Immunol 2020;11:633

24. Cornell CT, Kiosses WB, Harkins S, Whitton JL: Coxsackievirus B3 proteins directionally complement each other to downregulate surface major histocompatibility complex class I. J Virol 2007;81:6785-6797

25. Atkinson MA, Bowman MA, Campbell L, Darrow BL, Kaufman DL, Maclaren NK: Cellular immunity to a determinant common to glutamate decarboxylase and coxsackie virus in insulin-dependent diabetes. J Clin Invest 1994;94:2125-2129

26. Schloot NC, Willemen SJ, Duinkerken G, Drijfhout JW, de Vries RR, Roep BO: Molecular mimicry in type 1 diabetes mellitus revisited: T-cell clones to GAD65 peptides with sequence homology to Coxsackie or proinsulin peptides do not crossreact with homologous counterpart. Hum Immunol 2001;62:299-309

27. Kahan SM, Wherry EJ, Zajac AJ: T cell exhaustion during persistent viral infections. Virology 2015;479-480:180-193

28. McLane LM, Abdel-Hakeem MS, Wherry EJ: CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Ann Rev Immunol 2019;37:457-495

29. Viskari HR, Koskela P, Lonnrot M, Luonuansuu S, Reunanen A, Baer M, Hyoty H: Can enterovirus infections explain the increasing incidence of type 1 diabetes? Diabetes Care 2000;23:414-416

30. Nathanson N, Kew OM: From emergence to eradication: the epidemiology of poliomyelitis deconstructed. Am J Epidemiol 2010;172:1213-1229

31. Zinkernagel RM: Maternal antibodies, childhood infections, and autoimmune diseases. N Engl J Med 2001;345:1331-1335

32. Coppieters KT, Dotta F, Amirian N, Campbell PD, Kay TW, Atkinson MA, Roep BO, von Herrath MG: Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients. J Exp Med 2012;209:51-60

33. Feuer R, Mena I, Pagarigan R, Slifka MK, Whitton JL: Cell cycle status affects coxsackievirus replication, persistence, and reactivation in vitro. J Virol 2002;76:4430-4440

34. Saarinen NVV, Laiho JE, Richardson SJ, Zeissler M, Stone VM, Marjomäki V, Kantoluoto T, Horwitz MS, Sioofy-Khojine A, Honkimaa A: A novel rat CVB1-VP1 monoclonal antibody 3A6 detects a broad range of enteroviruses. Sci Rep 2018;8:1-13

35. Kirkegaard JS, Ravassard P, Ingvarsen S, Diedisheim M, Bricout-Neveu E, Gronborg M, Frogne T, Scharfmann R, Madsen OD, Rescan C, Albagli O: Xenotropic retrovirus Bxv1 in human pancreatic beta cell lines. J Clin Invest 2016;126:1109-1113

36. Weinzierl AO, Rudolf D, Maurer D, Wernet D, Rammensee HG, Stevanovic S, Klingel K: Identification of HLA-A*01- and HLA-A*02-restricted CD8+ T-cell epitopes shared among group B enteroviruses. J Gen Virol 2008;89:2090-2097

37. Mallone R, Mannering SI, Brooks-Worrell BM, Durinovic-Bello I, Cilio CM, Wong FS, Schloot NC: Isolation and preservation of peripheral blood mononuclear cells for analysis of islet antigen-reactive T cell responses: position statement of the T-Cell Workshop Committee of the Immunology of Diabetes Society. Clin Exp Immunol 2011;163:33-49

38. Oikarinen S, Tauriainen S, Hober D, Lucas B, Vazeou A, Sioofy-Khojine A, Bozas E, Muir P, Honkanen H, Ilonen J, Knip M, Keskinen P, Saha MT, Huhtala H, Stanway G, Bartsocas C, Ludvigsson J, Taylor K, Hyoty H, VirDiab Study G: Virus antibody survey in different European populations indicates risk association between coxsackievirus b1 and type 1 diabetes. Diabetes 2014;63:655-662

Article #2

Interferon-α skews the immunopeptidome of β cells towards the presentation of HLA-B ligands and promotes neo-antigen formation

Alexia Carré[1], Fatoumata Samassa[1], Annalisa Nicastri[2], Robert Parker[2], Javier Perez-Hernandez[1], Orlando Burgos Morales[1], Hanqing Liao[2], Maikel L. Colli[3], Decio L. Eizirik[3], Raphael Scharfmann[1], Piero Marchetti[4], Lorella Marselli[4], Sylvaine You[1], Nicola Ternette[2], Roberto Mallone[1,5]

[1]Université Paris Cité, Institut Cochin, CNRS, INSERM, 75014 Paris, France.
[2]The Jenner Institute, Nuffield Department of Medicine, University of Oxford, OX37DQ, UK.
[3]ULB Center for Diabetes Research, Université Libre de Bruxelles, 1070 Brussels, Belgium.
[4]Department of Clinical and Experimental Medicine, University of Pisa, 56126 Pisa, Italy.
[5]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, 75014 Paris, France

Corresponding author: Roberto Mallone, MD PhD - INSERM U1016 Cochin Institute, G.H. Cochin-Port-Royal, Bâtiment Cassini - 123, boulevard de Port-Royal – F-75014 Paris, France. Phone: +33-1-76.53.55.83 E-mail: roberto.mallone@inserm.fr.

ABSTRACT

Interferon (IFN)-α is the earliest cytokine signature observed in individuals at-risk of developing type 1 diabetes (T1D). Although IFN-α induces ER stress and HLA Class I (HLA-I) upregulation in β cells, its effect on the repertoire of HLA-I-bound peptides is unknown. Using immunopeptidomics, we characterized the differences in HLA-I peptide presentation between *in-vitro* resting and inflamed β-cell states. IFN-α increased peptide presentation, the diversity of source proteins and HLA-B-restricted ligands, while it decreased the presentation of peptides originating from secretory granule proteins. Neo-epitopes derived from alternative mRNA splicing, post-translational modifications and protein *cis*-splicing were instead predominantly presented in IFN-α-treated β cells. This antigenic landscape relied on processing by both the constitutive and immune proteasome. A subset of the β-cell peptides identified was recognized by circulating $CD8^+$ T cells with a variable effector/memory phenotype in healthy individuals, making them potential targets of islet autoimmunity. A cross-reactive PTPRN/MelanA epitope was recognized by high-frequency $CD8^+$ T cells. This comprehensive catalog of HLA-I-eluted peptides provides a key resource for characterizing diabetogenic $CD8^+$ T cells and developing T-cell-based biomarkers.

INTRODUCTION

Type I interferons (IFNs), notably IFN-α, play a crucial role in anti-viral immune responses by driving infected cells to express IFN-stimulated genes with anti-viral properties, thus limiting viral replication and spreading[328,329]. This response is further amplified by a feed-forward loop of IFN synthesis. Indeed, IFN-α release positively regulates the expression of IFN genes and regulatory factors[330,331], thus increasing its own production. Type I IFNs are also central in modulating innate immune responses and activating adaptive immunity[328,329]. They boost B-cell differentiation and antibody production[332,333], mature antigen-presenting cells, and induce the expression of human leukocyte antigen (HLA) class I (HLA-I) and class II (HLA-II) along with costimulatory molecules[334,335]. Upon detection of microbial components, macrophages and dendritic cells (DCs; mainly plasmacytoid DCs) secrete IFN-α, leading to additional cytokine and chemokine release. IFN-α also increases HLA-I expression directly on infected cells[301], thus augmenting their antigenic visibility, and enhance the cytolytic activity of NK and $CD8^+$ T cells and promote T-cell survival[336–339].

In front of their potent protective effect against viral infection, IFNs can also promote autoimmunity[328,340,341]. IFN immunotherapy to treat certain cancers and infections can induce autoimmune disorders[342–345], and monocytes exposed to the serum of lupus patients mature into DCs in an IFN-α-dependent manner[346]. Given the central role of DC maturation in switching from a tolerogenic to an immunogenic response, the resulting T-cell activation may prompt the autoimmune destruction of tissues. In type 1 diabetes (T1D), transcriptomics investigations on the peripheral blood of children genetically at risk revealed a strong type I IFN signature, preceding seroconversion[86,296]. Similarly, increased IFN-α levels are detected in the serum of T1D donors[347]. At the tissue level, IFN-stimulated genes are found overexpressed only in infiltrated islets of recent-onset T1D patients[348]. This observation is in line with the enhanced IFN-α expression specifically in pancreas specimens of T1D patients[51,87,349]. The predominance of IFN-α in the periphery and in the target tissue at early T1D stages may result from a viral infection. Indeed, one of the main environmental suspects in the onset of T1D autoimmunity is Coxsackievirus B (CVB)[350].

The most widely accepted histopathological hallmark of T1D is HLA-I hyperexpression in insulin-containing islets (ICIs)[142]. One early report documented that IFN-α is detected concomitantly to HLA-I overexpression in up to 93% of the ICIs in T1D patients[349]. More recently, a causal relationship between IFN-α, endoplasmic reticulum (ER) stress, and HLA-I upregulation has been demonstrated in β cells[299,301]. The increased surface HLA-I expression

on β cells may be a driver of T1D pathogenesis, as self-reactive $CD8^+$ T cells must recognize peptide-HLA-I (pHLA-I) complexes on β cells to trigger lysis. Additionally, both IFN-α and ER stress may impact the repertoire of HLA-I-presented peptides (so called immunopeptidome), as is the case for IFN-γ[351,352]. Indeed, ER stress promotes the generation of mRNA splice variants[299], which could result in presentation of neo-epitopes spanning novel splicing sites which may be targeted by $CD8^+$ T cells. Similarly, loss of β-cell ignorance could rely on other mechanisms of neo-epitope formation induced by cytokines[188], resulting from post-translation modifications (PTMs)[353] or peptide *cis*-splicing[239,303], i.e. the fusion of non-contiguous protein fragments. Inflammation also drives the assembly of the immunoproteasome[166,354], an enzymatic complex poised to degrade misfolded proteins, which might be responsible for the formation of *cis*-spliced peptides[178]. Moreover, β cells have highly secretory rates and might have the unique capability of sensitizing T cells at distance by releasing granules containing insulin (INS) and other granule proteins along with their byproducts[313]. By promoting ER stress, IFN-α may modulate the antigenic cargoes that are offered to T cells.

Therefore, IFN-α likely represents the missing link between CVB infection (and possibly other environmental triggers), inflammation, and early autoimmune activation in T1D[301,340]. Against this background, elucidating how IFN-α shapes the immunopeptidome of β cells, both in terms of antigen identity and processing pathways, is crucial to further understand the antigenic drivers of T1D. We have therefore used immunopeptidomics strategies applied to human β cells treated or not with IFN-α to define the antigenic display driving autoimmune T-cell recognition.

RESULTS

ECN90 β-cells exposed to IFN-α increase surface HLA-I expression and peptide presentation

To identify antigens possibly relevant to T1D pathogenesis, we cultured the ECN90 β-cell line[188] with or without IFN-α for 18 h. pHLA-I complexes were immunoprecipitated from cell lysates, and peptides sequenced by liquid chromatography tandem mass spectrometry (LC-MS/MS). ECN90 β cells exposed to IFN-α yielded more peptides than untreated cells (Fig. 1A-B), with respectively 91.1% and 74.9% sequences in the expected 8-14 amino-acid (aa) range for HLA-I ligands (Fig. 1A, E). This observation reflected the upregulation of surface HLA-I upon IFN-α exposure (Fig. 1C-D). Unsupervised Gibbs clustering of peptides uncovered distinct motifs characteristic of one or several HLA-I alleles expressed by ECN90 β cells. The predominant group 1 identified the HLA-A*03:01 (HLA-A3 from hereon) motif, group 2 corresponded to a mixture of HLA-A*02:01 (HLA-A2), -C*03:04 (-C3) and -C*07:01 (-C7) motifs and group 3 matched the similar motifs of HLA-B*40:01 (HLA-B40) and -B*49:01 (-B49) (Fig. S1).

Peptides were identified and filtered by a multi-step bioinformatics pipeline (Fig. 1E). The reference database used for the identification of genome-templated sequences was preliminarily complemented with translated mRNA isoforms enriched in human islets exposed or not to IFN-α and generating aa neo-sequences. Following selection for 8-14 aa length, and identification of mRNA variants, all sequences underwent filtering based on enriched expression of source proteins in β cells (see Methods). Peptides carrying PTMs were analyzed separately. All those spectra whose peptide interpretation did not match genome-templated, RNA-spliced or post-translationally modified sequences were searched for putative proteasomal *cis*-spliced peptides using the MARS algorithm[243]. In total, this workflow led to the identification of 4 mRNA splice candidates (Table S1), 784 conventional peptides (Table S2), 247 sequences carrying PTMs (Table S3), 11 *cis*-spliced sequences (Table S4). HLA-I restrictions were then predicted using NetMHCpan 4.1a[194]. For all peptide categories, more HLA-I binders were identified in cells exposed to IFN-α (Fig. 1E), in line with the induced surface HLA-I upregulation.

Collectively, these results show that IFN-α increases surface HLA-I expression and peptide presentation.

IFN-α increases the number and diversity of HLA-I-bound peptides and skews presentation toward HLA-B ligands

Next, we investigated the source proteins of the β-cell-enriched peptides identified. IFN-α treatment increased not only the number but also the diversity of source proteins, with 35.6% (58/163) of them appearing only in IFN-α-treated cells as compared to 6.1% (10/163) specific to untreated cells (Fig. S2A). Accordingly, 19 source proteins in IFN-α-treated cells yielded ≥2-fold more peptides (fold change $log_2FC>1$) than in the control condition (Fig. 2A, S2A). Conversely, only 8 source proteins displayed a similar $log_2FC<1$ in untreated vs. IFN-α-treated cells. In both conditions, the 5 source proteins yielding more peptides localize in secretory granules, i.e. chromogranin A (CHGA, n=132), INS (n=64), secretogranin 5 (SCG5, n=34); or are associated with them, being involved in their transport, i.e. microtubule-associated protein 1B (MAP1B, n=26), kinesin-like protein (KIF1A, n=34). Conversely, some known β-cell antigens were underrepresented (Fig. S2A), i.e. zinc transporter 8 (SLC30A8, n=5), glutamate decarboxylase (GAD2, n=3) and islet amyloid polypeptide (IAPP, n=1), or absent altogether, i.e. islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP) and INS defective ribosomal product (DRiP)[355]. Novel granule proteins were also identified, i.e. chromogranin B (CHGB, n=6), secretogranin 2 and 3 (SCG2, n=11; SCG3, n=10). Altogether, granule proteins contributed 52% of the immunopeptidome in untreated cells, which decreased to 31% upon IFN-α treatment (Fig. 2C), reflecting a dilutional effect due to the increased diversity of source proteins, and a smaller number (n=257 vs. 203 in basal vs. IFN-α-treated conditions) but equivalent abundance of granule peptides (Fig. 2D). The previously described preproINS$_{15-24}$ peptide[188] was the most abundantly presented in both basal and IFN-a-treated conditions.

To validate the results obtained on ECN90 β cells, HLA-I-bound peptides were identified from two HLA-A2/A3$^+$ primary human islet preparations from non-diabetic donors. Despite the limited amount of starting material and the dilution of β-cell-derived peptides with those from other endocrine cells, 375 β-cell-enriched conventional peptides were retrieved from these islet preparations, which is only 2-fold less than in ECN90 β cells (n=784). Across the 2 islet samples, the source proteins (n=72) of the peptides retrieved overlapped by 44% with those from ECN90 β cells (n=163) (Fig S2B). Interestingly, even though α cells account for only half the number of β cells in human islets (~30% vs. 60%)[356], the most represented protein was glucagon (n=42 peptides, Fig. S2C). Although islet samples contain several other cell types, the other dominant source proteins in islets were otherwise the same as in ECN90 β cells, notably CHGA (n=12) and INS (n=33; Fig. 2B). In total, 40/117 (34%) HLA-A2/-A3-restricted peptides identified in primary islets were also found in ECN90 β cells (Table S5), which was

only slightly lower than the overlap between the two islet preparations (56/117 shared peptides, 48%).

Another distinctive feature of IFN-α-treated ECN90 β cells was the significant enrichment in HLA-B binders. In untreated cells, only 8% of peptides were predicted HLA-B binders (-B40, -B49 or both), which increased to 34% in IFN-α-treated cells (Fig. 2E) and correlated to a 42-fold increase in gene expression (Fig. 2F). The relative representation of HLA-A, -C and -E ligands was instead not significantly different between the two conditions (41% vs. 40%, 5 vs. 9%, and 1% vs 1%, respectively), reflecting a more limited upregulation of their restriction molecules. This drastic HLA-B upregulation was even more surprising considering that it is the least expressed in basal conditions (data not shown). Thus, HLA-B ligands were the main contributor to the enriched peptide display induced by IFN-α. Interestingly, the dominant source proteins of HLA-B ligands were also different compared to HLA-A (Fig. S3). Up to 39% of HLA-B-restricted peptides originated from CHGA and INS in basal condition, compared to 7% for HLA-A ligands (Fig. 2G). Upon IFN-α exposure, the CHGA and INS dominance decreased 3.5-fold for HLA-B ligands compared to 1.5-fold for HLA-A. This again reflected their dilution into a more diversified immunopeptidome, to a much larger extent than for HLA-A ligands: HLA-B yielded 4.2-fold more source proteins (98 vs. 23 in IFN-α-treated vs untreated cells), compared to 1.2-fold for HLA-A ligands (107 vs. 89 in IFN-α-treated vs untreated cells) (Fig. S3). Using stringent criteria to assign restriction, 7 conventional peptides predicted to bind HLA-E*01:01 were also identified, none of which was specific to the basal condition, in line with the reported HLA-E upregulation induced by IFN-α[299]. A more extensive list of putative HLA-E ligands is provided in Table S6.

Collectively, these results show that IFN-α treatment increases not only the number of peptides displayed by HLA-I and HLA-E, but also their diversity in terms of source proteins. Granule-contained and granule-associated proteins are abundantly represented, with some novel ones (CHGB, SCG2, SCG3) identified. Moreover, IFN-α skews presentation toward other source proteins and HLA-B ligands.

The immunopeptidome of ECN90 β cells includes neo-epitopes generated by PTMs, mRNA splicing and peptide *cis*-splicing

We subsequently focused on candidate neo-epitopes generated by PTMs, mRNA splice variants or peptide *cis*-splicing. Thirty-four PTMs were identified on HLA-I-eluted ligands from ECN90 β cells (Table S7). Several of them (15/34, 44.1%) were found in the same proportion in both basal and IFN-α-treated samples (Fig. 3A). Since most PTMs are generated under cell stress

conditions[357], we elected to focus on those specifically found after IFN-α treatment (11/34, 32.4%). In order to not dismiss the opposite scenario of physiological PTMs that may be lost in an inflammatory environment, we also included those specific to the basal condition (8/34, 23.5%). The second selection criterion was the possibility to chemically introduce the modification on synthetic peptides for spectral matching purposes, excluding PTMs that naturally arise during peptide synthesis or MS acquisition, e.g. oxidation, deamidation, and all PTMs at the more vulnerable N- and C-terminal positions. Following these criteria, we retained for further study lysine acetylation, amidation, glutathione disulfide (S-glutathionylation), sulphone, cysteinylation, tryptophan conversion to kynurenin; and synthesized HLA-A and -B binders in both modified and unmodified form. Spectrum comparison with the modified peptides synthesized validated 12/18 (67%) sequences (Fig. 3B). In a second step, we compared the spectra of these validated sequences with their corresponding peptides synthesized in their native form. In this case, a spectral match would indicate an artifactual modification that is experimentally introduced. *Bona fine* biological PTMs were thus defined: cysteinylation (n=1 peptide), S-glutathionylation (n=8), phosphorylation (n=1) and sulphone (n=2) (Fig 3C). Interestingly, when considering all the identified glutathionylated sequences, 15 glutathionylated peptides were specifically found in untreated cells, derived mostly (12/15) from insulin and were non-binders, while those specific to IFN-α-treated cells mostly (9/10) originated from proteins other than insulin and were predicted HLA binders (9/10) (Table S3). The immunopeptidome of human islets did not retrieve glutathionylated peptides.

Despite being a relevant PTM in T1D[232,358,359], no citrullination was identified. However, citrullinated arginine has the same mass increase as deamidated glutamine or asparagine, and distinguishing them by MS is challenging[360]. Citrullination may thus be mis-identified as deamidation in peptides containing both an arginine and a glutamine or asparagine.

When analyzing the 4 mRNA splice candidates identified by spectral matching with the synthetic peptide, only one HLA-B40/B49-restricted INS-205 sequence (Fig. 3D) was validated and was exclusively found in IFN-α-treated cells (Table S1).

For *cis*-spliced peptides, most spectra were matched (9/11, 81.8%), with some isoleucine/ leucine ambiguities not resolvable by MS (Table S4).

Altogether, 12 peptides with PTMs, 1 mRNA splice and 9 *cis*-spliced peptides were validated by spectral matching.

The antigenic landscape of β cells depends on (immuno-)proteasomal activity

To explore whether the generation of the identified peptides was proteasome-dependent, we characterized changes in the immunopeptidome of ECN90 β cells exposed or not to IFN-α and (immuno-)proteasome inhibitors. The catalytic subunits of the constitutive and immuno-proteasome were found expressed by ECN90 β cells, with the latter upregulated by IFN-α treatment as expected (Fig. S4A). Treatment with carfilzomib, an inhibitor of the β5 proteasome subunit, or ONX-0914, a specific inhibitor of the β5i immunoproteasome subunit with minimal cross-reactivity for its conventional β5 counterpart, led to a significantly decreased chymotrypsin-like activity (Fig. 4A). The HLA-I-eluted ligands were filtered and processed following the previous bioinformatics pipeline, focusing on conventional, post-translationally-modified and *cis*-spliced candidates. As observed before, HLA-B-ligands contributed the most to the significant increase in HLA-I-restricted peptides in IFN-α-treated cells. This feature was conserved upon treatment with proteasome and immuno-proteasome inhibitors (Fig. S4B). When clustering the β-cell-enriched peptides retrieved across conditions, several distinctive groups were uncovered (Fig. 4B). Some peptides remained unaffected by the inhibition of the constitutive or the immuno-proteasome in untreated (group 3) and IFN-α exposed cells (group 8). Conversely, peptides in group 2, specific to resting cells, were dependent on proteasomal activity, while peptides in groups 6 and 9, specific to IFN-α-treated cells, were dependent on constitutive and/or immuno-proteasome catalysis. Interestingly, the expression of certain peptides was only enriched in cells treated with carfilzomib, regardless of IFN-α exposure (group 1) or only in resting cells (group 4). Peptides enriched upon inhibition of the immuno-proteasome were also observed (group 5), including some shared with cells exposed to both IFN-α and proteasome inhibitor (group 7). Of note, the INS_{15-24} peptide was not affected by (immuno-)proteasome inhibition and was consistently identified in all the conditions. Only one cis-splice candidate was found (RIINEPTAA/K, from HSPA5 also known as GRP78) in IFN-α and ONX-0914-treated cells.

Collectively, these results show the expression and enzymatic activity of the constitutive and immuno-proteasome in β cells, and their role in shaping the β-cell immunopeptidome.

HLA-A2-restricted peptides are recognized by circulating $CD8^+$ T cells

We previously documented that the vast majority of individuals harbor similar frequencies of circulating, largely naïve $CD8^+$ T cells reactive to islet epitopes[119,188,260]. Hence, the existence of a cognate naïve $CD8^+$ T-cell repertoire is the first requirement to qualify a given peptide as epitope. We therefore selected 72 HLA-A2-restricted peptides to evaluate their recognition by

circulating $CD8^+$ T cells in HLA-$A2^+$ healthy donors, using combinatorial double-coded HLA-A2 multimers (MMrs) loaded with the corresponding synthetic peptides[119,188,260]. This panel included peptides derived from granule proteins (n=12), granule-associated KIF1A (n=3), a selection of other β-cell-enriched proteins (n=42), S-glutathionylated peptides (n=5), $CD8^+$ T-cell epitopes described in the literature[245] and not identified in our dataset (n=10) and control viral (CMV, Flu) and MelanA epitopes (n=3). A representative staining is shown in Fig. S5 and results are detailed in Table S8. We retained as epitopes those peptides with a cognate $CD8^+$ T-cell frequency falling into the expected range of naïve T-cell precursors (1-50/10^6 $CD8^+$ T cells)[361] or above. Out of the 72 peptides tested, 28 (39%) were validated as recognized with the above frequency cut-off in at least 2 donors (Fig. 5A). Notably, validated epitopes included the granule-derived $ABCC8_{394-402}$, $SCG5_{186-195}$ previously described[188], granule-associated $KIF1A_{419-427}$, novel peptides derived from the known β-cell antigen IA-2/PTPRN ($PTPRN_{57-66}$, $PTPRN_{526-534}$) and C33-glutathionylated $PCSK1N_{24-33}$. Only a fraction of reference epitopes from the literature were found recognized: $GAD2_{141-149}$, INS_{6-14} and $PTPRN_{962-970}$. As reported before[119,188], most epitopes were recognized by T cells harboring a limited effector/memory fraction in most donors ($CD45RA^+CCR7^-$, $CD45RA^-CCR7^-$, or $CD45RA^-CCR7^+$; Fig. 5B), even when frequencies above the naïve precursor range were observed (i.e. $ELAV3_{156-164}$, $SCG5_{186-195}$, $PTPRN_{962-970}$, C33-glutathionylated $PCSK1N_{24-33}$). Exceptions were noted from some novel epitopes targeted by predominantly effector/memory T cells, namely $CADPS_{680-688}$, $HSPA5_{2-10}$ and $PEG10_{36-44}$. Conversely, Flu $MP_{58–66}$-reactive $CD8^+$ T cells included as viral positive controls displayed the expected predominantly effector/memory phenotype, which was instead more heterogenous for CMV $pp65_{495-503}$.

One surprising observation was the absence of positive control $CD8^+$ T cells recognizing $MelanA_{26-35}$ epitope, which is targeted by highly immunodominant naïve precursors in most HLA-$A2^+$ individuals[119]. We therefore asked whether this could reflect cross-reactivity with an islet epitope included in the same staining panel. Indeed, the combinatorial analysis of double-labeled MMr^+ cells is based on the selection of those T cells labeled with the 2 fluorochromes coding for the corresponding peptides, and by the exclusion of T cells labeled with more than 2 fluorochromes. The binding of MMrs loaded with 2 different peptides would therefore result in the exclusion of these cross-reactive cells. We therefore searched for such cross-reactivity by looking at all other peptides sharing one fluorochrome (PE-CF594 or BV785) with $MelanA_{26-35}$[232]. We found that $PTPRN_{830-839}$ (PE- and BV785-labeled) also scored negative due to the sharing of the BV785 fluorochrome with $MelanA_{26-35}$. Indeed, when including triple-

labeled T cells (PE/PE-CF594/BV785 MMr^{+}), a fully PTPRN/MelanA cross-reactive population was readily visualized (Fig. 5C).

Collectively, these results show that several β-cell peptides identified are recognized by circulating $CD8^{+}$ T cells with a variable effector/memory phenotype in healthy individuals, making them potential targets of islet autoimmunity. Such recognition may be favored by high naïve precursor frequencies and by cross-reactivity with other unrelated epitopes.

DISCUSSION

Our study provides an in-depth view of the repertoire of HLA-I-bound peptides presented by β cells exposed or not to the early cytokine IFN-α. Expectedly, IFN-α exposure increased HLA-I surface expression[299,301], which resulted into a higher number and abundance of presented peptides. These peptides originated from a wider array of source proteins, possibly reflecting IFN-α-induced ER stress[301] and increased numbers of mis-folded proteins undergoing proteasomal degradation[178]. Within this more diversified peptide display of IFN-α-treated β cells, fewer peptides derived from secretory granule proteins were retrieved, although the abundance of the remaining granule peptides was unaffected. Apart from a dilutional effect, this likely reflects the inhibitory effect of IFN-α on the synthesis of insulin[362] and other granule proteins. Alternatively, this effect could result from the reported β-cell dedifferentiation induced by inflammation, which downregulates several β-cell-specific genes[363].

On one hand, β-cell dedifferentiation may represent a defense mechanism to limit autoimmune vulnerability[279,364], and the parallel increase in HLA-E-presented peptides that we observed may also be at play. HLA-E inhibits natural killer cell-mediated cytotoxicity[134] and reportedly presents a limited set of peptides to regulatory $CD8^+$ T cells[365,366]. On the other hand, IFN-α treatment preferentially upregulated HLA-B, as reported for IFN-γ[351,367], and skewed peptide presentation toward HLA-B ligands, which were the main contributors to the overall HLA-I increase. Mechanistically, this reflects the structure of the HLA-B gene locus, which harbors two IFN response elements, while HLA-A and -C loci have only one[351,367,368]. Moreover, HLA-B ligands originated from a very restricted set of source proteins, with CHGA and INS accounting for 39% and 11% of the total number of peptides displayed under basal and IFN-α-treated conditions as compared to 7% and 4% for HLA-A. This implies that the inflammatory microenvironment of insulitis may skew the autoimmune response toward a different set of peptides and recruit additional $CD8^+$ T-cell clonotypes. This finding is also relevant in view of the paucity of HLA-B-restricted β-cell epitopes described in the literature[245]. The novel HLA-B40- and -B49-restricted peptides that we identified may provide a relevant resource to assign antigen specificities to the orphan T-cell receptors of islet-infiltrating $CD8^+$ T cells of T1D patients[369].

The immunopeptidome obtained confirmed our previous finding[188] that some known β-cell antigens are underrepresented (ZnT8, GAD65, IAPP) or even absent (IGRP, INS DRiP). It is of note that this was the case despite a much deeper peptide sequencing depth compared to our previous work[188]. Using higher cell numbers and improved though more restrictive bioinformatics pipelines, we retrieved a total of 714 8-12mer conventional peptides as

compared to the previous 78. This does not exclude that these missing antigens may become T-cell targets upon peptide presentation by antigen-presenting cells uptaking β-cell material. The INS DRiP case is noteworthy, as it has been confirmed as a relevant T-cell antigen targeted by islet-infiltrating $CD8^+$ T cells[369]. The short-lived and unstable nature of DRiPs may explain this discrepancy, emphasizing the need for combined antigen identification strategies[254] using ribosome profiling[370] and T-cell receptor fingerprinting[371]. We also identified novel antigens, namely CHGB, SCG2 and SCG3, which share several features with the previously reported SCG5, UCN3 and PCSK2[188,260]. They are all soluble granule proteins undergoing intermediate processing by proconvertases and furins to yield their bioactive products, which are released with secretory granules. This processing, along with the modulation of their biosynthesis according to metabolic demand and inflammatory context, may easily divert them toward the HLA-I pathway under β-cell stress conditions[178].

The comparison of the immunopeptidome of ECN90 and primary human islet cells was also informative. Several peptides restricted for the shared HLA-A2/A3 alleles identified in the ECN90 β-cell line were also found in islets, lending support to this model. Intriguingly, several glucagon-derived peptides were also retrieved from islets (n=42 vs. 33 for insulin). This finding suggests that the autoimmune resistance of α cells cannot be ascribed to a lesser antigen presentation leading to immune ignorance. Other α-cell-intrinsic defense mechanisms and/or a higher proneness of the T-cell repertoire to recognize INS rather than glucagon[369] may be at play.

Our thorough investigation of HLA-eluted peptide neo-sequences led to the robust identification of 1 HLA-B-restricted INS-205 mRNA splice peptide; 9 *cis*-spliced candidates and 12 sequences bearing PTMs. Both the INS-205 and the majority (6/9) of the *cis*-spliced peptides were found exclusively after IFN-α treatment, underlining the role of IFN-α in shaping the β-cell neo-antigen landscape. While decreased presentation of INS and other granule peptides may reduce the antigenic visibility of β cells, the generation of neo-antigens and the preferential upregulation of HLA-B may increase it and contribute to epitope spreading. Of note, one *cis*-spliced candidate was also correctly assigned to a reported alternative mRNA splice isoform, supporting the reliability of the MARS algorithm[243]. The most represented PTM (8/12) was S-glutathionylation, which involves the formation of a disulfide bond between a cysteine residue and a glutathione tripeptide. Glutathione biosynthesis is catalyzed by γ-glutamyl-cysteine synthetase, whose expression is induced by insulin[372]. Glutathione is a key circulating anti-oxidant, whose levels are reduced in T1D patients due to increased utilization[373,374]. It participates to insulin degradation by reducing the disulfide bonds between

the A and B chain[375]. The observation that less and different glutathionylated peptides are found in IFN-α-treated conditions likely reflects the altered redox balance of inflamed β cells. Moreover, epitope glutathionylation can alter T-cell recognition[376], implying that oxidative stress may also contribute to the antigenic visibility of β cells. We validated the recognition of the glutathionylated peptide $PCSK1N_{24-33}$ by circulating $CD8^+$ T cells, which were among the top-scoring in terms of frequency. This recognition was specific to the glutathionylated version as no cognate T cells were retrieved for the native form, confirming that PTMs may play a key role in T-cell recognition. Cysteines are exquisitely sensitive to oxidative PTMs, with cysteinylation (i.e. a disulfide-linked cysteine addition) also found in our dataset and previously reported in HLA-I-restricted T-cell epitopes[377,378].

Using (immuno-)proteasome inhibitors, a peculiar pattern of proteasome-dependent and independent peptide presentation was observed. This pattern likely reflects the presence or absence of protein motifs preferred by the proteasome chymotrypsin-like activity targeted by these inhibitors. Lack of proteasome dependence may reflect the reliance of these peptides on the trypsin- and caspase-like activities of the proteasome, or processing in other cell compartments. One notable example was the INS_{15-24} peptide, which is reportedly processed by signal peptidases during the import of nascent proINS into the ER[379]. The other relevant instance was the *cis*-spliced peptide which may originate from the reported islet epitope GRP78[231,232]. It was exclusively found in β cells exposed to immuno-proteasome inhibitor, suggesting that its formation follows other pathways.

Peptides presented at the β-cell surface are relevant to the pathogenesis of T1D only if there are $CD8^+$ T cells able to recognize them, indicating their participation in autoimmune processes. We have started to address this question by screening healthy donors for a subset of the HLA-A2-restricted peptides identified, and validated several of them as targets of circulating T cells. Their frequency fell mostly in the range of naïve precursors previously reported for other islet epitopes[119,188,260]. Two novel observations are however noteworthy. First, some of them were recognized at frequencies above this range, despite limited effector/memory fractions of their cognate T cells. This observation underscores the importance of the starting naïve precursor frequency in shaping the autoimmune repertoire. Second, an unexpectedly predominant effector/memory phenotype was noted for other islet epitopes, suggesting the possibility of cross-reactivity with other unrelated foreign epitopes. Indeed, we identified one instance of such cross-reactivity for a PTPRN/MelanA peptide, which resulted in high T-cell frequencies. The role that epitope mimicry mechanisms may play in triggering autoimmunity is increasingly recognized[62,380–382] and warrants further investigation. Cross-reactivity between two different

islet epitopes is another possibility, which may explain the lack of T-cell recognition for some epitopes previously described. Overall, the large repertoire of candidate epitopes identified by immunopeptidomics will require higher-throughput technologies based on DNA-barcoded MMrs[383] for a comprehensive validation of those targeted by $CD8^+$ T cells. One key question will be to define whether such technologies hold a suitable sensitivity to detect $CD8^+$ T cells that fall mostly in the naïve frequency range.

In conclusion, our study shows that IFN-α shapes the immunopeptidome of β cells, drastically increasing the presentation of HLA-B-restricted and neo-antigenic peptides and decreasing the presentation of secretory granule peptides. Circulating $CD8^+$ T cells recognize at least a subset of these peptides, and this recognition may be shaped by the naïve T-cell repertoire and by cross-reactivity with other epitopes. This comprehensive catalog of HLA-I-eluted peptides invites further studies to identify diabetogenic $CD8^+$ T cells and develop T-cell-based biomarkers.

ACKNOWLEDGMENTS. We gratefully acknowledge L. Bailly, myBrain Technologies, for his help in developing the bioinformatics analysis pipeline; S. Salignac, Cochin Institute, for technical assistance; and the Cochin Institute CYBIO Flow Cytometry Core Facility for assistance with cell analysis.

FUNDING. This work was supported by The Leona M. and Harry B. Helmsley Charitable Trust (1901-03689), Agence Nationale de la Recherche (ANR-19-CE15-0014-01) and Fondation pour la Recherche Medicale (EQU20193007831), to R.M.; by the European Foundation for the Study of Diabetes (EFSD/JDRF/Lilly European Programme in Type 1 Diabetes Research 2019), to N.T. R.M. and D.L.E. received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust.

DUALITY OF INTEREST. No potential conflicts of interest relevant to this article were reported.

METHODS

β-cell culture and treatment and flow cytometry

The ECN90 β-cell line[188] was maintained in DMEM/F12 Advanced medium (ThermoFisher) supplemented with 2% bovine serum albumin (BSA; fraction V, fatty-acid-free; Roche), 50 µM β-mercaptoethanol (Sigma), 10 mM nicotinamide (Merck), 1.7 ng/mL sodium selenite (Sigma), 100 U/mL penicillin-streptomycin (Gibco), 1% GlutaMAX (Gibco). Cells were seeded at 8-8.5×10^6 cells in 75 cm² or 16.5×10^6 cells in 150 cm² culture flasks (TPP) coated with 0.25% fibronectin from human plasma (Sigma) and 1% extracellular matrix from Engelbreth-Holm-Swarm murine sarcoma (Sigma) and cultured at 37°C in 5% CO_2 for 18-24h. ECN90 cells were treated in DMEM/F12 medium (ThermoFisher), supplemented as above or without BSA with the following agents and final concentrations: IFN-α (PBL Assay Science, #11100-1; 2,000 U/mL), carfilzomib (Selleckchem, #S2853; 50 nM), ONX-0914 (Selleckchem, #S7172; 100 nM) and forskolin (Tocris; 10 µM).

Surface HLA-I expression was verified on ECN90 β cells stained for 30 min with Live/Dead Violet (ThermoFisher) and monoclonal antibody (mAb) to HLA-A,B,C (RRID:AB_2917755) before acquisition on an LSRFortessa and analysis with FlowJo v10.8.

Human pancreatic islets were obtained from 2 non-diabetic brain-dead organ donors (79-year-old male, BMI 27.8 kg/m²; 80-year-old female, BMI 21.6 kg/m²; protocol approved by the Ethics Committee of the University of Pisa, Italy).

Western blotting and measurement of proteasome enzymatic activity

Cell pellets were lysed in RIPA buffer with protease inhibitor cocktail (Sigma) and total protein was quantified with Pierce BCA protein assay kit (ThermoFisher). Protein samples were separated by Bolt 4–12% Bis-Tris polyacrylamide gels (Invitrogen) under denaturing conditions. Detection of HLA-A,B,C was performed using the HC10 mAb (RRID:AB_2728622; 1/1000) followed by a horseradish peroxidase-conjugated horse anti-mouse IgG (RRID:AB_330924; 1/1000). α-tubulin mAb was used for loading control (RRID:AB_1210456; 1/1000).

ECN90 cells were collected by trypsinization, extensively washed with phosphate-buffered saline (PBS) and the enzymatic activity of the proteasome/immunoproteasome subunits measured using Proteasome-Glo (Promega #G1180).

Real-time quantitative PCR (RT-qPCR)

Total RNA was isolated from ECN90 cells using RNeasy mini kit (Qiagen) and RNA concentration and purity was assessed by Nanodrop (ThermoFisher). cDNA was synthesized from 500 ng total RNA using superscript VILO synthesis kit (ThermoFisher #11754050) prior to gene expression analysis of proteasome/immunoproteasome subunits by real-time quantitative PCR (LightCycler, Roche). Quantitect SYBR green PCR mastermix (Qiagen) along with primers designed with Primer3 software were used (Table S9). For HLA-I genes, primers were as reported[351]. Melting curve analysis was performed to evaluate the specificity of each amplicon prior to analysis. To normalize gene expression across samples, *ACTB* (β-actin) and *GAPDH* were used as housekeeping genes. The ΔCt comparative method for relative quantification was used.

HLA immunoprecipitation and peptide elution

Purified antibodies (8-16 mg; produced in-house) were incubated with protein A Sepharose beads for 30 min before being washed with borate buffer (0.05 M boric acid, 0.05 M KCl, 4 mM NaOH, pH 8.0). Bound antibodies were cross-linked to the beads with 40mM dimethyl pimelimidate dihydrochoride (Sigma) in borate buffer pH 8.3 for 30 min. The cross-linking was terminated with ice-cold 0.2 M Tris pH 8.0, and unbound antibodies removed by washing with 0.1 M citrate buffer pH 3.0 followed by 50 mM Tris pH 8.0.

Dry-frozen cell pellets were lysed in 1 mL lysis buffer (1% IGEPAL-CA 630, 300 mM NaCl, 100 mM Tris pH 8.0, 1X Roche cOmplete Mini protease inhibitor Cocktail EDTA-free) per 1×10^8 ECN90 cellsby mixing at 4°C for 30 min. Human primary islets were lysed in lysis buffer in bead beater tubes (Precellys Evolution, Bertin) following 5 cycles at 7,200 rpm for 20s, separated by 20 s pauses. Lysates were collected, further mixed for 20 min at 4°C, and cleared by centrifugation at 500g for 10 min followed by 21,000g for 45 min at 4°C. Peptide-HLA complexes were captured by incubating the cleared lysates with anti-HLA-I W6/32-cross-linked Sepharose A beads overnight at 4°C. The resin was collected by gravity.

HPLC fractionation of ECN90 HLA-I-bound peptides

Samples were resuspended in 120 µL loading buffer (0.1% TFA, 1% acetonitrile) and injected in an Ultimate 3000 HPLC System (ThermoFisher). Peptides were separated across a 4.6- by 50-mm ProSwift RP1S column (ThermoFisher) using 1mL/min flow rate over 10 min from 2% to 35% buffer B (0.1% TFA in acetonitrile) in buffer A (0.1 TFA in water). Fifteen fractions

were collected, every 30 s, and peptide fractions 1-9 were combined into odd and even aliquots and vacuum-dried prior to LC-MS/MS acquisition.

Purification of HLA-I-eluted peptides

Resuspended HLA-I-eluted samples (1:2 and 1:1, respectively) were centrifuged through 5 kDa cutoff filters (Merck Millipore #UFC3LCCNB-HMT) and vacuum-dried. They were then resuspended in loading buffer and cleared using Pierce C18 Spin Tips (ThermoFisher, 84850). Final elution was performed in 30% acetonitrile 0.1% TFA for HLA-I-bound peptides and 35% acetonitrile 0.1% TFA for HLA-II-bound peptides. Samples were vacuum-dried prior to LC-MS/MS acquisition.

LC-MS/MS acquisition

Dried samples of HLA-I-bound peptides from ECN90 cells exposed or not to IFN-α and synthetic peptides were resuspended in loading buffer (1% acetonitrile, 0.1% TFA) and analyzed by an Ultimate 3000 RSLCnano system coupled to a Fusion Lumos mass spectrometer (ThermoFisher). Peptides were separated using a PepMap C18 column, 75µm x 50cm, 2µm particle size (ThermoFisher) with a 30 min (synthetic peptides and proteasomal digests) to 60 min linear acetonitrile in water gradient of 2-25%, at a flow rate of 250 µL/min. Solvent contained 5% DMSO and 0.1% formic acid (v/v). Peptides were ionized using an EasySpray source at 2,000 V and ions were introduced into the mass spectrometer through an on-transfer tube at 305°C. Data-dependent acquisition was performed with one full MS1 spectra recorded from 300 to 1,500 m/z (120,000 resolution, 400,000 AGC target, 60 ms accumulation time), followed by MS2 scans (30,000 resolution, 120 ms accumulation, 300,000 AGC target). Precursor selection was performed using TopSpeed mode at a cycle time of 2 s. High-collision dissociation (HCD) fragmentation was induced at an energy setting of 32 for singly charged peptides and of 28 for peptides with a charge state 2-4.

Alternatively, vacuum-dried samples of HLA-I-bound peptides from ECN90 cells exposed to (immuno-)proteasome inhibitors or human islets were resuspended in SCP loading buffer (1% acetonitrile, 0.1% formic acid) and analyzed by a nanoElute system coupled to a timsTOF SCP mass spectrometer (Bruker). Peptides were loaded into an Aurora C18 column, 25cm x 75µm, 1.7µm particle size (IonOpticks) with a loading pressure of 800 bars for 9 min and were separated with a linear acetonitrile gradient of 2-25% in 0.1% acetic acid over 60 min and 24-37% for additional 6 min at a flow rate of 150 nL/min at 50°C. Peptides were ionized with the CaptiveSpray source (Bruker) at 1,400V and 180°C. Data were acquired in DDA PASEF mode

with one TIMS-MS survey and 10 PASEF MS2 scans per cycle. Ion accumulation and ramp time in the TIMS analyzer were set to 166ms each. The ion mobility range for peptide analysis was set to 1/K0 = 1.7 to 0.7 Vs cm² and the m/z range was 100-1,700. Two compound regions were defined using m/z and 1/K0 vertex's as follows (300, 0.7; 800, 1.2; 800, 0.9; 500, 0.7) and (700, 1.4, 700, 1.1; 1000, 1.7; 1500, 1.7). Precursors with charge states 1-3 and a minimum threshold of 500 arbitrary units (a.u.) were fragmented and re-sequenced until reaching a "target value" of 20,000 a.u. Collision energies were 70 eV at 1/K0 = 1.7 Vs cm-2; 40 eV at 1/K0 = 1.34 Vs cm-2; 40 eV at 1/K0 = 1.1 VS Vs cm-2; 30 eV at 1/K0 = 1.06 Vs cm-2; 20 eV at 1/K0 = 0.7 Vs cm-2.

LC-MS/MS data analysis

Analysis of the raw data was performed using Peaks X or Peaks X Pro (Bioinformatics Solutions), reporting 5-10 peptides for each spectrum in *de-novo* sequencing, and searching a protein sequence FASTA file containing the reviewed human Uniprot entries (downloaded on 22/01/2019), predicted peptide neo-sequences translated from published RNA sequencing datasets[317] and previously reported defective ribosomal products[355]. Peaks PTM searches were performed with all 314 built-in modifications. Some specific PTM searches were run by setting the PTM of interest as a variable modification in both the De Novo and the DB search. The false discovery rate (FDR) was calculated with a decoy database search integrated in Peaks and set to 1% for all samples. A 5% FDR was allowed when looking for mRNA variants reported in human islets (see *Immunopeptidomics bioinformatics analysis*). Data analysis was performed with Python and Excel. Sequence clustering was generated by GibbsCluster (https://services.healthtech.dtu.dk/). Venn diagrams were graphed using the online tool https://bioinformatics.psb.ugent.be/webtools/Venn/ and venn2 in Python. All other figures and statistical analyses were done using GraphPad Prism (v9.0).

RNAseq datasets

RNAs from 6 individual preparations of primary human islets exposed or not to IFN-α for 18 h[299] and from HLA-II^{lo} and HLA-II^{hi} human mTECs[188] were sequenced on an Illumina HizSeq 2000 at high depth (coverage >150×10^6 reads, which is sufficient to detect >80% slice variants). Gene expression was quantified using Salmon version 0.13.2[384] with parameters "--seqBias –gcBias --validateMappings." GENCODE v31 (GRCh38)[385] was chosen as the reference genome and has been indexed with the default k-mer values. The estimated number of reads obtained from Salmon were used as input to perform differential expression with DESeq2

1.24.0[384]. For each gene included in DESeq2's model, a log2 fold change (FC) was computed and a Wald test statistic was assessed with a *p* value and an adjusted *p* value. Transcripts were considered differentially expressed when presenting a FC >1.50 and an adjusted *p* value <0.05. Only transcripts presenting >0.5 transcripts per million (TPM) in at least 20% of samples were selected for further analysis. Datasets have been deposited under GEO: GSE148058.

Transcriptomics bioinformatics analysis

For the RNAseq pipeline, 30,947 mRNAs (TPM >0.5) were filtered based on:

a) A median TPM >5 in islets, either under basal or inflammatory conditions, a cut-off selected based on the median TPM of known islet antigens (islet expression filter; n=12,594).

b) A median TPM <0.1 in mTECs (either HLA-II^{lo} or HLA-II^{hi}), or a median TPM fold-decrease >100 vs. islet (mTEC expression filter).

c) A median TPM fold-increase >10 in islets compared to 12 control tissues (adipose tissue, breast, colon, heart, kidney, liver, lung, lymph node, ovary, prostate, skeletal muscle, white blood cells), using the Illumina BodyMap 2.0 dataset (islet enrichment filter). Tissues of neuroendocrine origin (brain, testis, adrenal gland and thyroid) were excluded for this filtering.

d) We subsequently focused our analysis on mRNA isoforms, as described[188]. The predicted translation products were aligned using the R package Biostrings v2.52.0, with BLOSUM100 as the score matrix, and aa neo-sequences were defined by comparing the predicted aa sequence of each mRNA isoform with that of the reference (canonical) mRNA, taking as reference the longest and/or most prevalent mRNA isoform in islets (neo-sequence generation filter).

Immunopeptidomics bioinformatics analysis

An in-house analysis pipeline was designed (Python 3.7) and applied to the Peaks PTM output file. Long peptides (>14 amino acids long) were filtered out. Remaining sequences were sorted based on their source proteins/genes and according to their expression in the pancreas, based on downloadable databases and datasets. Briefly, peptides whose source genes were found in the pancreas at the RNA level, or at the protein level with a high/medium evidence according to the Human Protein Atlas (V18) or the Human Protein Reference Database (release 9) were kept. Previously published single-cell RNA sequencing datasets from human pancreatic islets[386] were re-processed to generate a list of genes that are enriched in endocrine cells. Peptides identified in MS were further retained only if their source gene was found in this list. Together, these steps defined the "β-cell-enriched expression" filter (Fig. S2). In parallel, the spectra found in the Peaks PTM search (matched to the database) were excluded from the Peaks

DeNovo search output, and the remaining spectra were searched for *cis*-spliced peptide sequences in ECN90 β-cell preparation. The 10 reported peptides per spectrum were investigated for putative forward (i.e. ligation of two fragments in the order in which they occur in the parent protein) or backward (i.e. ligation of two fragments in the reverse order compared to the parent protein sequence) *cis*-spliced sequences of proteins included in the reference database. Sequences were then fed into the MARS v1.0 algorithm[243], trained with the genome-templated sequences. The MARS output and the putative *cis*-spliced sequences were compared, and sequences found in both were retained and filtered based on their β-cell-enriched expression as above. NetMHCpan4.1a (https://services.healthtech.dtu.dk/) was installed locally and used to predict HLA-binders to the alleles expressed by the cell line, i.e. HLA-A*02:01, -A*03:01, -B*40:01, -B*49:01, -C*03:04, -C*07:01 and -E*01:01 (cutoff score of <2)[194]. For post-translationally-modified peptides, the prediction was performed on native sequence. HLA-I binding predictions were not computed for *cis*-spliced sequences, as netMHCpan ranking is a selection parameter already embedded in the MARS algorithm. A single HLA prediction was attributed to a peptide if the score difference with the second-best allele was ≥3-fold. Otherwise, the two best HLA prediction were assigned. An HLA-E restriction was assigned only in the absence of other predicted restrictions.

HLA-A2 multimer staining

The recognition of 72 HLA-A2-restricted candidate epitopes and 3 control epitopes by circulating $CD8^+$ T cells was assessed with our previously described combinatorial dual-labeling HLA-A2 multimer (MMr) assays[119,188,232,260],. As the assay setup allows for the simultaneous testing of 15 peptides in a unique PBMC sample, 5 panels of 15 epitopes/each were individually tested on 5 PBMC aliquots from the same draw of 5 healthy individuals (median age 36 yrs, range 29-49X-Y; 3 females, 2 males). A median of 1.1×10^6 $CD8^+$ T cells (preliminarily enriched by magnetic negative selection) were acquired, and T-cell frequencies are reported only when ≥3 MMr^+ events were counted to limit undersampling bias. The effector/memory phenotype is reported only if ≥5 MMr^+ cells were acquired.

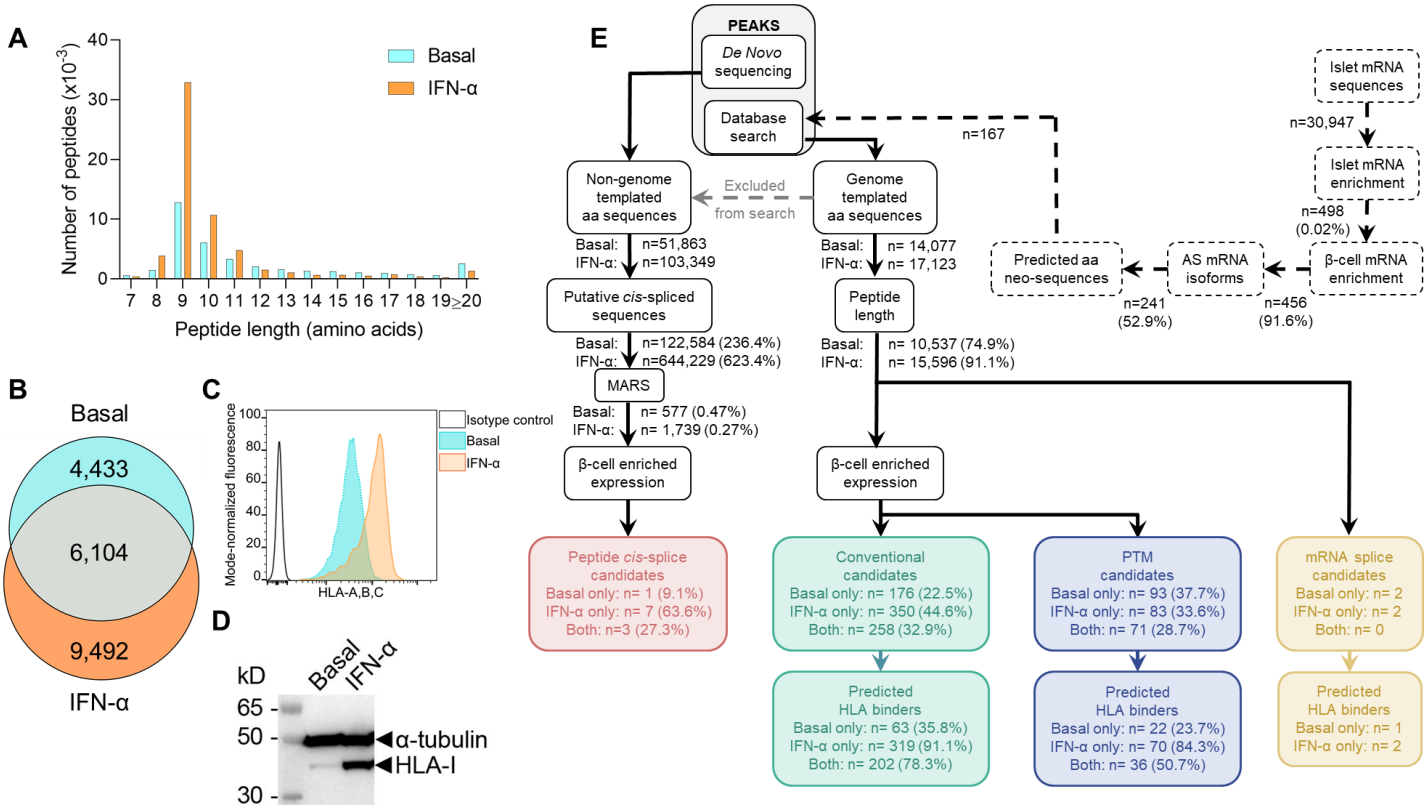


Figure 1. Immunopeptidome profiling of HLA-I-bound peptides presented by ECN90 β cells. A. Length distribution of HLA-I-eluted peptides from ECN90 β cells under basal (blue) and IFN-α-stimulated conditions (orange). Four biological replicates (3.5-4.5×10^8 cells/each) were acquired for each condition, and unique peptides across replicates were counted. **B.** Number of 8-14mer peptides eluted from the above conditions. **C.** HLA-I expression detected by flow cytometry on ECN90 β cells in basal (blue) and IFN-α-stimulated conditions (orange). **D.** HLA-I heavy chain expression detected by Western blot in whole cell lysates of ECN90 cells exposed or not to IFN-α, with α-tubulin bands as loading controls. **E.** Bioinformatics analysis pipeline. Predicted aa neo-sequences from RNAseq datasets of human islets (dashed lines) were appended to the database used for immunopeptidome search. Database-matched sequences identified by PEAKS (grey box) were sequentially filtered based on their length, on whether they match mRNA variants (yellow box) and on the enriched expression of their source proteins in β cells. Conventional candidates (green box) and sequences carrying PTMs (blue box) were separated. HLA-I-binding predictions were performed using NetMHCpan4.1a. In parallel, non-genome-templated peptides were interpreted as potential *cis*-spliced candidates and fed into the MARS algorithm, followed by filtering according to the enrichment of their source proteins in β cells (red box).

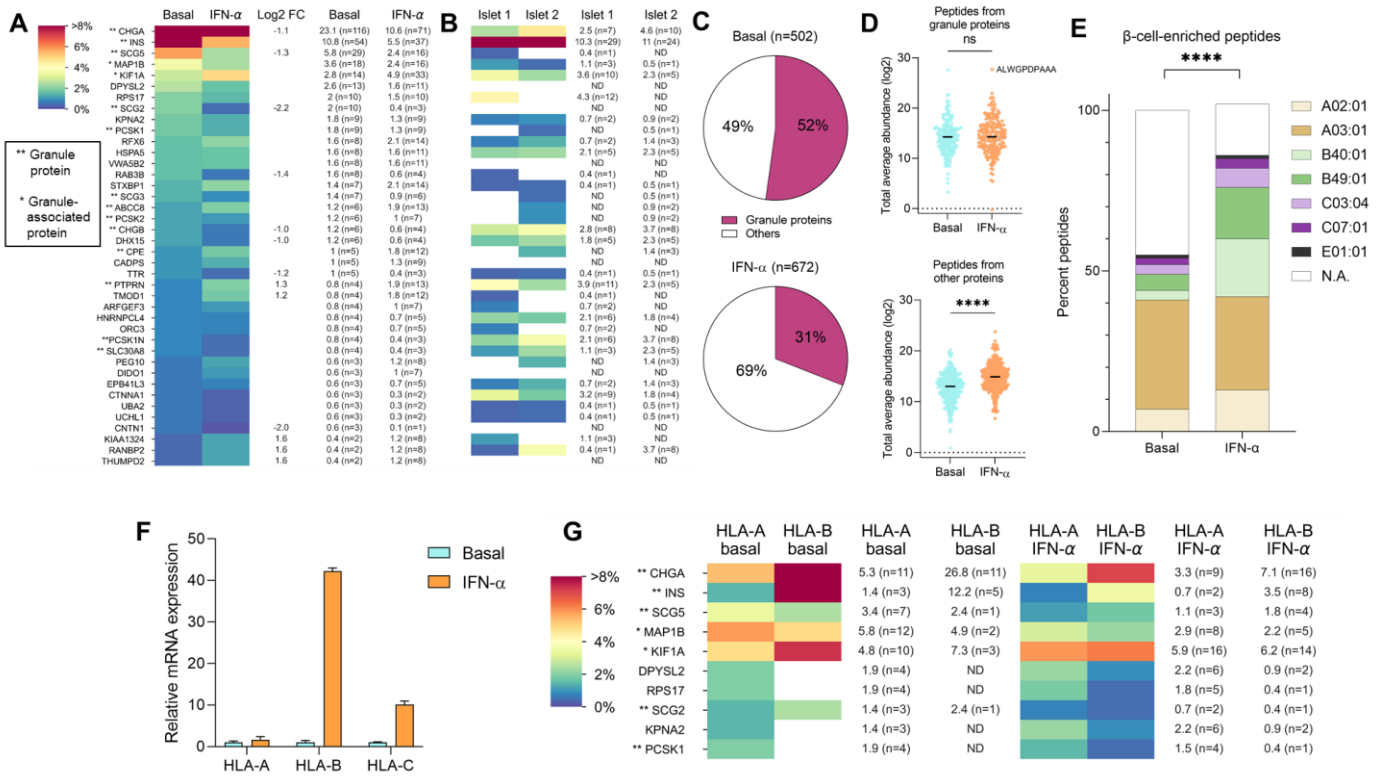


Figure 2. Comparison of the immunopeptidome of ECN90 β cells exposed or not to IFN-α. A-B. Heatmap of the relative representation of the top 40 source proteins in ECN90 β cells (A) and human islets (B). The color scale is proportional to the number of ligands identified for each protein out of the total number of ligands in a given condition, expressed as percentage. Only conventional and post-translationally-modified sequences are included. Peptides carrying PTMs were counted as such only for PTMs defined as likely biological, they were otherwise counted as unmodified; *cis*-spiced peptides were excluded. Percent values and peptide numbers are listed on the right. Source proteins enriched in IFN-α-treated cells (log2 fold change, FC ≥ 1) or in basal condition (log2 FC ≤ 1) are indicated for ECN90 β cells. **C.** Proportion of peptides originating from granule proteins in basal (top) and IFN-α-treated conditions (bottom). Granule-associated proteins are not included. **D.** Average total abundance of conventional peptides originating from granule (top) and other proteins (bottom). The peptides were identified by Peaks and quantified by Progenesis. The INS_{15-24} sequence of the most abundant peptide identified in both conditions is indicated. Horizontal bar represents the median value. ****p<0.0001 by Wilcoxon test. **E.** Relative distribution of predicted HLA-I ligands for each allele expressed by ECN90 β cells between conditions. ****P<0.0001 by χ^2 test. **F.** Relative mRNA expression of classical HLA-I alleles in ECN90 β cells exposed or not to IFN-α. GAPDH was used as an internal normalizing control, and each gene was normalized to the basal sample. Data represent mean ± SEM of 3 biological replicates. **G.** Heatmap of the relative representation of the top 10 source proteins in ECN90 β cells based on predicted HLA-A or HLA-B restriction.

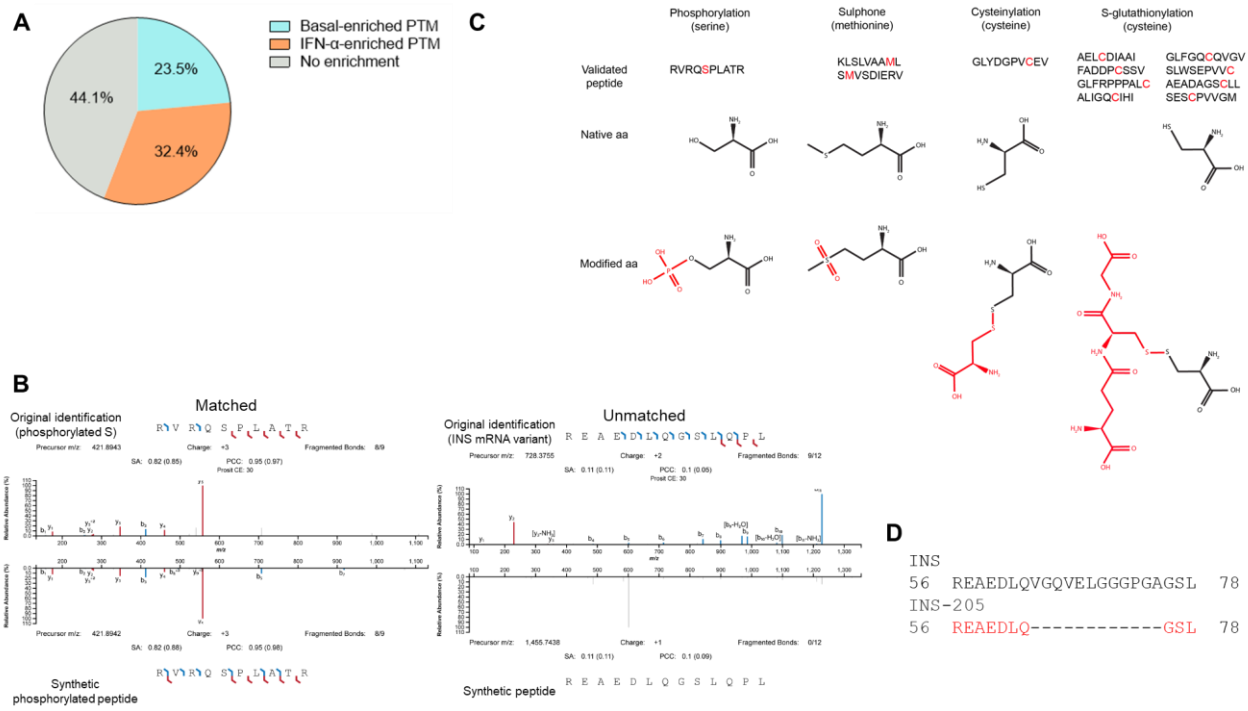


Figure 3. Validation of HLA-I-eluted candidate neo-epitopes. A. Global PTM enrichment in basal and IFN-α-treated condition. **B.** Spectral matching examples of matched (left) and unmatched (right) synthetic peptides compared to the initial peptide identification using an online tool (https://www.proteomicsdb.org/use). **C.** Validated post-translationally modified peptides and representation of the native and modified aa. The PTM is indicated in red. **D.** Peptide alignment of INS and INS-205 mRNA variant. The sequence of the HLA-eluted variant is indicated in red and correspond to a sequence spliced out as compared to the canonical INS mRNA.

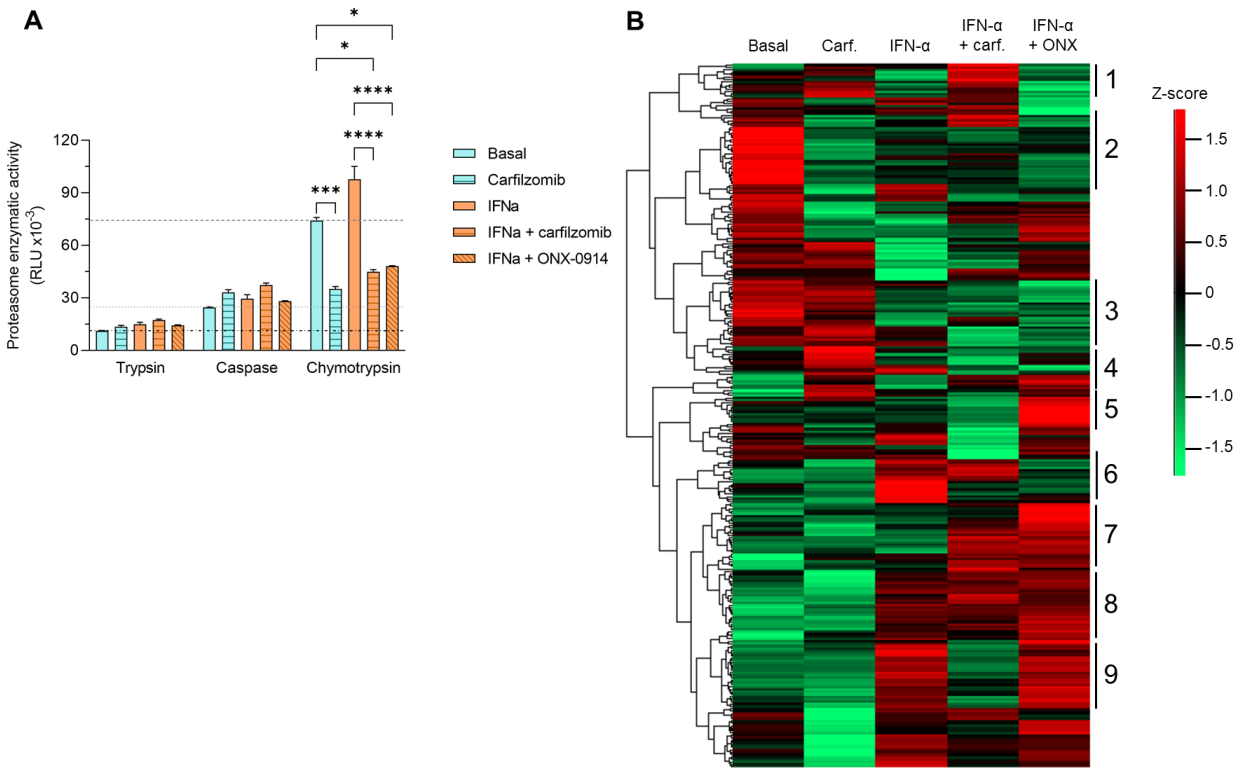


Figure 4. Inhibition of constitutive and immuno-proteasome on ECN90 β cells. A. Resting and IFN-α-exposed ECN90 β cells (3-5×10^6, n=3 biological replicates) were treated with carfilzomib or ONX-0914 and enzymatic activity of all three catalytic subunits were measured. Data represent mean ± SEM of triplicates. The horizontal lines indicate the basal activities. *P<0.01, *** P<0.0005, ****P<0.0001 by two-way ANOVA test. **B.** For each precursor found in ≥2 replicates in one of the 5 conditions, ion intensities were summed. For peptides identified on several instances, median values were computed. The values were log2-transformed, missing values were imputed from normal distribution downshifted by 1.8 and normalization was computed using row z-scores. The heatmap was generated with Perseus (v2.0.6), clustering rows with "average" as the agglomeration method and "Pearson correlation" as the distance matrix. Vertical bars indicate distinct peptide clusters (see Results).

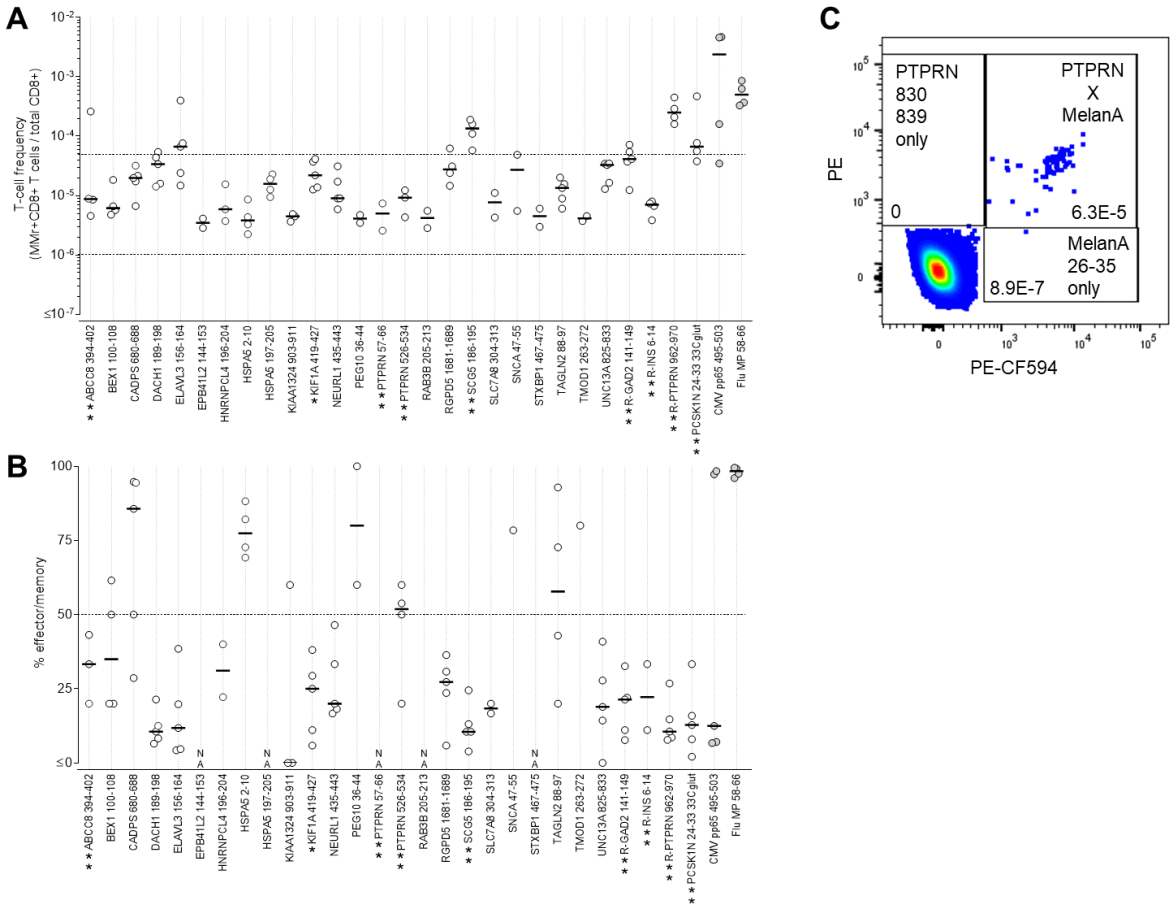


Figure 5. Circulating CD8$^+$ T cells recognize candidate HLA-A2-restricted islet epitopes. **A.** Frequencies of MMr$^+$CD8$^+$ out of total CD8$^+$ T cells reactive to candidate HLA-A2-restricted islet epitopes in healthy HLA-A2$^+$ adults (n=5). Only epitopes scoring positive in ≥2 donors with ≥3 MMr$^+$ cells counted are displayed (complete list in Table S8). At least 4.8×10^5 (median 1.1×10^6) CD8$^+$ T cells were acquired. Viral control epitopes CMV pp65$_{495\text{-}503}$ and Flu MP$_{58\text{-}66}$ are shown in gray. Horizontal bars indicate the expected frequency range of naïve CD8$^+$ T cells. **: granule protein, *: granule-associated protein. **B.** Percent effector/memory out of total MMr$^+$CD8$^+$ T cells. Phenotype was not assigned (NA) when <5 MMr$^+$CD8$^+$ T cells were counted. Bars indicate median values of each distribution. **C.** The gating strategy based on positivity for 2 MMr fluorochromes precludes visualization of cross-reactive CD8$^+$ T cells recognizing two different epitopes, since these T cells are positive for 3 fluorochromes. Given the lack of MelanA$_{26\text{-}35}$ MMr$^+$CD8$^+$ T cells, possible cross-reactivities were analyzed for all islet epitopes included in the same staining panel. Cross-reactive PTPRN$_{830\text{-}839}$ MMr$^+$ cells (labeled with PE and BV785) and MelanA$_{26\text{-}35}$ MMr$^+$ cells (labeled with PE-CF594 and BV785) were thus identified by gating cells positive for only 2 fluorochromes (PE/BV785$^+$, PE-CF594$^-$, labeled as "PTPRN$_{830\text{-}839}$ only"; or PE-CF594/BV785$^+$, PE$^-$, labeled as "MelanA$_{26\text{-}35}$ only") or for all 3 (PE/PE-CF594/BV785$^+$, labeled as "PTPRN × MelanA"). The resulting overlaid dot plot is displayed.

Table S1. HLA-I-eluted peptides derived from mRNA splice variants identified in ECN90 β cells. The first column displays the peptide sequence with, in parenthesis, the mass shift of PTMs found on the aa. PTMs are detailed in the next column and were ultimately defined as likely artifactual. Subsequent columns indicate the condition and the number of replicates in which the peptides were identified; aa length; UniProt accession number of the canonical source protein; gene name and description; whether the source is a granule protein (Y, yes; N, no); predicted HLA-I restrictions and ranks based on NetMHC4.1a (NA, not assigned). The last column indicates whether the peptide sequence was validated by spectral matching (ND, not determined).

Peptide	PTM		Presence	Basal - # of replicates	IFN-α - # of replicates	Length	UniProt	Gene name	Gene description	Granule protein	HLA restriction	Rank	Spectral matching
REAEDLQGSL		0	IFN-α	0/4	1/4	10	P01308	mRNA_INS-205	Insulin	Y	B40:01/B49:01	0.043/0.347	Yes
REAEDLQGSLQPL		0	IFN-α	0/4	1/4	13	P01308	mRNA_INS-205	Insulin	Y	B40:01	0.838	No
SLDSYPVTSK		0	Basal	1/4	0/4	10	Q9UM54	mRNA_MYO6-208	Unconventional myosin VI	N	A03:01	0.0701	No
M(+42.01)(+15.99)DRGEQGLLRTD	Acetylation (Protein N-term); Oxidation(M)		Basal	2/4	0/4	12	P55327	mRNA_TPD52-209	Tumor protein D52	N	NA	NA	ND

Table S2. Conventional HLA-I-eluted peptides identified in ECN90 β cells. The first column displays the peptide sequence. Subsequent columns indicate the condition and the number of replicates in which the peptides were identified; aa length; UniProt accession number; aa position within the source protein; gene name and description; whether the source is a granule protein (Y, yes; N, no). Predicted HLA-I restrictions based on NetMHC4.1a and the corresponding ranks are shown in the two last columns (NA, not assigned).

Peptide	Presence	Basal - # of replicates	IFN-α - # of replicates	Length	UniProt	Position	Gene name	Gene synonyms	Gene description	Granule protein	HLA restriction	Rank
VEVNVIRV	IFN-α	0/4	4/4	8	Q09428	185-192	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B49:01	0.125
KIYNKIMHL	IFN-α	0/4	1/4	9	Q09428	394-402	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A02:01	0.0403
KEADFSPSV	Both	1/4	2/4	9	Q09428	567-575	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B49:01	0.0045
KEADFSPSVAF	IFN-α	0/4	4/4	11	Q09428	567-577	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B40:01	0.1182
SEFLSSAEI	Basal	1/4	0/4	9	Q09428	611-619	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B49:01	0.0116
IIFESPFNK	Both	3/4	4/4	9	Q09428	787-795	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A03:01	0.01
HLSDHLMQA	IFN-α	0/4	1/4	9	Q09428	862-870	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A02:01	0.0564
GTIQREGTLK	Both	2/4	4/4	10	Q09428	905-914	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A03:01	0.3892
VEWTGLKV	IFN-α	0/4	1/4	8	Q09428	1085-1092	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B49:01	0.1383
YEARFQQKL	IFN-α	0/4	1/4	9	Q09428	1218-1226	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B40:01	0.0117
VLIAAVTSI	IFN-α	0/4	1/4	9	Q09428	1259-1267	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A02:01	0.2575
GLLAPSLIPK	Basal	2/4	0/4	10	Q09428	1327-1336	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A03:01	0.0647
LLAPSLIPK	Both	1/4	1/4	9	Q09428	1328-1336	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	A03:01	0.0235
KRGAILEF	IFN-α	0/4	1/4	8	Q09428	1552-1559	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	C07:01	0.3578
LEFDKPEKL	IFN-α	0/4	4/4	9	Q09428	1557-1565	ABCC8	ABC36,HHF1,HI,I	ATP binding cassette subfamily C member 8	Y	B40:01/B49:01	0.0175/0.0366
TETDLLNGHL	IFN-α	0/4	2/4	10	P45844	44-53	ABCG1	ABC8	ATP binding cassette subfamily G member 1	N	B40:01	0.2829
TEAQRFSSL	IFN-α	0/4	1/4	9	P45844	61-69	ABCG1	ABC8	ATP binding cassette subfamily G member 1	N	B40:01/B49:01	0.0872/0.1167
KRLAIALEL	IFN-α	0/4	1/4	9	P45844;Q9H172	223-231	ABCG1;ABCG4	ABC8,WHITE2	ATP binding cassette subfamily G member 1;AT	N	C07:01	0.0305
GEYGDQNSRL	IFN-α	0/4	4/4	10	P45844	331-340	ABCG1	ABC8	ATP binding cassette subfamily G member 1	N	B40:01	0.0417
GEYGDQNSRLV	IFN-α	0/4	4/4	11	P45844	331-341	ABCG1	ABC8	ATP binding cassette subfamily G member 1	N	B49:01	0.1368
AALGTMTSL	IFN-α	0/4	2/4	9	P45844	538-546	ABCG1	ABC8	ATP binding cassette subfamily G member 1	N	C03:04	0.0372
GEYIAPEKI	Both	3/4	4/4	9	Q9ULC5;P33121	550-558	ACSL1;ACSL5	ACS1,ACS2,ACS5	Acyl;CoA synthetase long;chain family member	N	B49:01	0.008
REFFNDLDL	IFN-α	0/4	2/4	9	P51826	1186-1194	AFF3	LAF4,MLLT2-like	AF4/FMR2 family member 3	N	B40:01	0.1072
ATDPTIHTV	IFN-α	0/4	1/4	9	Q9H6U8	450-458	ALG9	DIBD1	ALG9, alpha-1,2-mannosyltransferase	N	A02:01	0.0376
KVRDQQLVY	Basal	1/4	0/4	9	Q9UJX6	802-810	ANAPC2	APC2,KIAA1406	Anaphase promoting complex subunit 2	N	A03:01/C07:01	0.1741/0.3096
SLWGSLKPK	Both	2/4	4/4	9	O00213	695-703	APBB1	Fe65,RIR	Amyloid beta precursor protein binding family B	N	A03:01	0.0335
GIYYVPKGK	Both	1/4	4/4	9	Q70E73;Q7Z5R6	424-432	APBB1IP;RAPI	ALS2CR18,ALS2C	Amyloid beta precursor protein binding family B	N	A03:01	0.0514
SLYHRVLLY	Both	2/4	4/4	9	Q5TH69	345-353	ARFGEF3	BIG3,C6orf92,KI/	ARFGEF family member 3	N	A03:01	0.0084
SLAMPGEVK	Basal	1/4	0/4	9	Q5TH69	1163-1171	ARFGEF3	BIG3,C6orf92,dJ	ARFGEF family member 3	N	A03:01	0.3328
FIHDILTEV	Both	4/4	4/4	9	Q5TH69	1232-1240	ARFGEF3	BIG3,C6orf92,KI/	ARFGEF family member 3	N	A02:01	0.0038
KMPLKPIFL	Both	1/4	2/4	9	Q5TH69	1407-1415	ARFGEF3	BIG3,C6orf92,KI/	ARFGEF family member 3	N	E01:01	0.0042
GEEKTIQV	IFN-α	0/4	3/4	8	Q5TH69	1743-1750	ARFGEF3	BIG3,C6orf92,dJ	ARFGEF family member 3	N	B49:01	0.3581
REFDTSPGL	IFN-α	0/4	2/4	9	Q5TH69	1782-1790	ARFGEF3	BIG3,C6orf92,dJ	ARFGEF family member 3	N	B40:01/B49:01	0.007/0.0164
TISKLMTEYKK	IFN-α	0/4	1/4	11	Q5TH69	2021-2031	ARFGEF3	BIG3,C6orf92,dJ	ARFGEF family member 3	N	A03:01	1.8555
REIDRQEVI	IFN-α	0/4	2/4	9	O15085	731-739	ARHGEF11	GTRAP48,KIAA0	Rho guanine nucleotide exchange factor 11	N	B49:01	0.0097
EEGPIIKEI	IFN-α	0/4	4/4	9	O15085	801-809	ARHGEF11	GTRAP48,KIAA0	Rho guanine nucleotide exchange factor 11	N	B49:01	0.1306
RMLERVIGR	Both	1/4	3/4	9	Q9Y2G3	1135-1143	ATP11B	ATPIF,ATPIR,KIA	ATPase phospholipid transporting 11B (putative	N	A03:01	0.0649
SEEGKSLWEL	IFN-α	0/4	3/4	10	Q93084	43-52	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	B40:01	0.1907
KEYEPEMGKV	IFN-α	0/4	4/4	10	Q93084;P16615	120-129	ATP2A3;ATP2	SERCA3,ATP2B,I	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	B49:01	0.046
KEYEPEMGKVI	IFN-α	0/4	3/4	11	Q93084	120-130	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	B49:01	0.0809
YEPEMGKVI	IFN-α	0/4	3/4	9	Q93084	122-130	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	B49:01/B40:01	0.2008/0.3741
IEIKSTTL	IFN-α	0/4	2/4	8	Q93084	166-173	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	B40:01/B49:01	0.242/0.3345
VATGLHTEL	IFN-α	0/4	2/4	9	Q93084	224-232	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	C03:04	0.0066
MRKEFTLEF	Both	1/4	3/4	9	Q93084	479-487	ATP2A3	SERCA3	ATPase sarcoplasmic;endoplasmic reticulum Ca	N	C07:01	0.0045
IAMGSGTAV	IFN-α	0/4	1/4	9	Q93084;P16615	718-726	ATP2A3;ATP2	SERCA3,ATP2B,I	ATPase sarcoplasmic/endoplasmic reticulum Ca	N	C03:04	0.0881
IEDPVRPEV	Both	2/4	4/4	9	P23634;Q16720;P	670-678	ATP2B4;ATP2	ATP2B2,CFAP39,	ATPase plasma membrane Ca2; transporting 3;,	N	B49:01	0.0216
GTNDGPALKK	IFN-α	0/4	4/4	10	P23634;P20020;Q	786-795	ATP2B4;ATP2	ATP2B2,MXRA1,	ATPase plasma membrane Ca2+ transporting 4;	N	A03:01	0.0437
ISAIPTRSL	IFN-α	0/4	4/4	9	P23634	1046-1054	ATP2B4	ATP2B2,MXRA1,	ATPase plasma membrane Ca2+ transporting 4	N	C03:04	0.0288
SIQKPYNQK	IFN-α	0/4	1/4	9	P23634	1153-1161	ATP2B4	ATP2B2,MXRA1,	ATPase plasma membrane Ca2+ transporting 4	N	A03:01	0.0368
RVRQPILQY	IFN-α	0/4	1/4	9	Q9HBH7;Q9BXY8	54-62	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-lir	N	A03:01	0.0402
QLMEKLREK	IFN-α	0/4	2/4	9	Q9HBH7;Q9BXY8	91-99	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-lir	N	A03:01	0.0744

Table S2, continued

KLREKQLSH	IFN-α	0/4	4/4	9	Q9HBH7;Q9BXY8	95-103	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-lir	N		A03:01	0.1979
REKQLSHSL	IFN-α	0/4	4/4	9	Q9HBH7;Q9BXY8	97-105	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-lir	N		B40:01/B49:01	0.0292/0.0869
ASNSVILHL	Both	2/4	4/4	9	O75973;Q86Z23	216-224	C1QL1;C1QL4	C1QRF,C1QTNF1	Complement C1q like 1;Complement C1q like 4	N		C07:01	0.2929
IEDPNFKNKV	IFN-α	0/4	4/4	10	Q9NY47	192-201	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit	N		B49:01	0.23
KRLTPNYTL	IFN-α	0/4	3/4	9	Q9NY47	535-543	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit	N		C07:01	0.0411
REPVTLDFL	IFN-α	0/4	3/4	9	Q9NY47	572-580	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit	N		B40:01/B49:01	0.4353/0.88
FESEGHVFI	IFN-α	0/4	4/4	9	Q9NY47	678-686	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit	N		B49:01	0.0199
RSLDNHGYVFK	IFN-α	0/4	1/4	11	Q9NY47	792-802	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit	N		A03:01	0.3304
VEEESGKEVL	IFN-α	0/4	3/4	10	Q9ULU8	15-24	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		B40:01	0.1288
MEVEGGEKL	IFN-α	0/4	2/4	9	Q9ULU8	423-431	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		B40:01	0.0065
VILHPTPNSPK	Basal	1/4	0/4	11	Q9ULU8	480-490	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		A03:01	0.3391
SLFEMVQRL	Both	1/4	3/4	9	Q9ULU8	680-688	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		A02:01	0.0064
YLRDLLERA	IFN-α	0/4	1/4	9	Q9ULU8	730-738	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		A02:01	0.2579
SLLERVLMK	Both	3/4	4/4	9	Q9ULU8;Q86UW7	815-823	CADPS2;CADF	CAPS,CAPS1,CAP	Calcium dependent secretion activator;Calcium	N		A03:01	0.0095
SSFLSFTVK	Both	3/4	1/4	9	Q9ULU8	1206-1214	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		A03:01	0.2016
KTYETIRNR	Both	4/4	4/4	9	Q9ULU8	1313-1321	CADPS	CAPS,CAPS1,KIA/	Calcium dependent secretion activator	N		A03:01	0.0269
ISEEGFHYL	IFN-α	0/4	3/4	9	Q13555;Q13557;(	80-88	CAMK2G;CAM	CAMKG,CAMKD,	Calcium/calmodulin dependent protein kinase I	N		C03:04	0.0974
KANLPLLQR	IFN-α	0/4	1/4	9	O43439;Q06455;(	179-187	CBFA2T2;RUN	MTGR1,ZMYND3	CBFA2/RUNX1 translocation partner 2;RUNX1 t	N		A03:01	0.3124
SRPEFIKHL	Both	2/4	4/4	9	Q49MI3	464-472	CERKL	RP26	Ceramide kinase like	N		C07:01	0.0111
RSAAVLALL	Both	2/4	3/4	9	P10645	2-10	CHGA		Chromogranin A	Y		E01:01	0.2083
VEVISDTL	IFN-α	0/4	1/4	8	P10645	37-44	CHGA		Chromogranin A	Y		B40:01/B49:01	0.357/0.8383
RGDERILS	IFN-α	0/4	1/4	8	P10645	61-68	CHGA		Chromogranin A	Y		NA	NA
SILRHQNLLK	Both	3/4	4/4	10	P10645	68-77	CHGA		Chromogranin A	Y		A03:01	0.1217
NLLKELQDL	Both	3/4	2/4	9	P10645	74-82	CHGA		Chromogranin A	Y		A02:01	0.6193
LLKELQDLAL	Both	1/4	3/4	10	P10645	75-84	CHGA		Chromogranin A	Y		A02:01	1.5166
LLKELQDLALQG	IFN-α	0/4	2/4	12	P10645	75-86	CHGA		Chromogranin A	Y		NA	NA
LKELQDLALQG	IFN-α	0/4	2/4	11	P10645	76-86	CHGA		Chromogranin A	Y		NA	NA
KELQDLAL	IFN-α	0/4	4/4	8	P10645	77-84	CHGA		Chromogranin A	Y		B40:01/B49:01	0.4181/1.154
HSGFEDELSE	Basal	1/4	0/4	10	P10645	97-106	CHGA		Chromogranin A	Y		NA	NA
HSGFEDELSEV	Both	1/4	1/4	11	P10645	97-107	CHGA		Chromogranin A	Y		NA	NA
HSGFEDELSEVL	Both	1/4	3/4	12	P10645	97-108	CHGA		Chromogranin A	Y		B40:01	0.6446
HSGFEDELSEVLE	Both	4/4	4/4	13	P10645	97-109	CHGA		Chromogranin A	Y		NA	NA
HSGFEDELSEVLEN	Both	3/4	2/4	14	P10645	97-110	CHGA		Chromogranin A	Y		NA	NA
SEVLENQSSQAEL	IFN-α	0/4	4/4	13	P10645	105-117	CHGA		Chromogranin A	Y		B40:01	0.7643
VLENQSSQAELKE	Basal	1/4	0/4	13	P10645	107-119	CHGA		Chromogranin A	Y		NA	NA
LENQSSQAELKE	Basal	1/4	0/4	12	P10645	108-119	CHGA		Chromogranin A	Y		NA	NA
NQSSQAELKE	Basal	1/4	0/4	10	P10645	110-119	CHGA		Chromogranin A	Y		NA	NA
EAVEEPSSKDVME	Basal	1/4	0/4	13	P10645	119-131	CHGA		Chromogranin A	Y		NA	NA
AVEEPSSKD	Basal	1/4	0/4	9	P10645	120-128	CHGA		Chromogranin A	Y		NA	NA
AVEEPSSKDV	Basal	1/4	0/4	10	P10645	120-129	CHGA		Chromogranin A	Y		B49:01	1.6371
AVEEPSSKDVM	Basal	1/4	0/4	11	P10645	120-130	CHGA		Chromogranin A	Y		B40:01	1.6169
AVEEPSSKDVME	Both	1/4	1/4	12	P10645	120-131	CHGA		Chromogranin A	Y		NA	NA
VEEPSSKDV	Both	2/4	4/4	9	P10645	121-129	CHGA		Chromogranin A	Y		B49:01	0.2372
VEEPSSKDVM	IFN-α	0/4	4/4	10	P10645	121-130	CHGA		Chromogranin A	Y		B40:01/B49:01	0.5364/1.5396
VEEPSSKDVME	Basal	1/4	0/4	11	P10645	121-131	CHGA		Chromogranin A	Y		NA	NA
ALPEPMQESK	Both	4/4	4/4	10	P10645	153-162	CHGA		Chromogranin A	Y		A03:01	0.3446
PEPMQESK	Basal	1/4	0/4	8	P10645	155-162	CHGA		Chromogranin A	Y		NA	NA
GLSAEPGWQA	Basal	1/4	0/4	10	P10645	216-225	CHGA		Chromogranin A	Y		A02:01	0.8389
VVLNPHPSLGY	Both	2/4	2/4	11	P10645	252-262	CHGA		Chromogranin A	Y		NA	NA
VVLNPHPSLGYK	Both	3/4	1/4	12	P10645	252-263	CHGA		Chromogranin A	Y		A03:01	0.5553
VLNPHPSLGY	Both	4/4	4/4	10	P10645	253-262	CHGA		Chromogranin A	Y		A03:01	0.2125
KEIRKGESRSEAL	Both	3/4	4/4	13	P10645	263-275	CHGA		Chromogranin A	Y		B40:01	1.3685

Table S2, continued

EIRKGESRSEAL	Basal	1/4	0/4	12	P10645	264-275	CHGA		Chromogranin A	Y	NA	NA
GESRSEAL	IFN-α	0/4	4/4	8	P10645	268-275	CHGA		Chromogranin A	Y	B40:01	0.4029
AVVPQGLFR	Basal	1/4	0/4	9	P10645	310-318	CHGA		Chromogranin A	Y	A03:01	0.3147
AVVPQGLFRGGK	Both	4/4	4/4	12	P10645	310-321	CHGA		Chromogranin A	Y	A03:01	0.355
RGGKSGELEQEEE	Basal	1/4	0/4	13	P10645	318-330	CHGA		Chromogranin A	Y	NA	NA
RGGKSGELEQEEER	Basal	1/4	0/4	14	P10645	318-331	CHGA		Chromogranin A	Y	NA	NA
GKSGELEQEEERLS	Basal	1/4	0/4	14	P10645	320-333	CHGA		Chromogranin A	Y	NA	NA
KSGELEQEEERLS	Basal	1/4	0/4	13	P10645	321-333	CHGA		Chromogranin A	Y	NA	NA
SGELEQEEERL	Basal	1/4	0/4	11	P10645	322-332	CHGA		Chromogranin A	Y	NA	NA
SGELEQEEERLS	Basal	1/4	0/4	12	P10645	322-333	CHGA		Chromogranin A	Y	NA	NA
GELEQEEERL	IFN-α	0/4	3/4	10	P10645	323-332	CHGA		Chromogranin A	Y	B40:01	0.1737
GELEQEEERLS	Both	1/4	1/4	11	P10645	323-333	CHGA		Chromogranin A	Y	NA	NA
ELEQEEERLS	Basal	1/4	0/4	10	P10645	324-333	CHGA		Chromogranin A	Y	NA	NA
LEQEEERLS	Basal	1/4	0/4	9	P10645	325-333	CHGA		Chromogranin A	Y	NA	NA
KEWEDSKRW	Basal	1/4	0/4	9	P10645	334-342	CHGA		Chromogranin A	Y	B49:01/B40:01	0.5071/1.181
WSKMDQLAKE	Basal	1/4	0/4	10	P10645	342-351	CHGA		Chromogranin A	Y	NA	NA
WSKMDQLAKELTAE	Both	2/4	1/4	14	P10645	342-355	CHGA		Chromogranin A	Y	NA	NA
SKMDQLAKELT	Basal	1/4	0/4	11	P10645	343-353	CHGA		Chromogranin A	Y	NA	NA
SKMDQLAKELTAE	Both	4/4	4/4	13	P10645	343-355	CHGA		Chromogranin A	Y	NA	NA
KMDQLAKE	Both	1/4	1/4	8	P10645	344-351	CHGA		Chromogranin A	Y	NA	NA
KMDQLAKEL	Both	4/4	4/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406
KMDQLAKELT	Both	4/4	4/4	10	P10645	344-353	CHGA		Chromogranin A	Y	NA	NA
KMDQLAKELTA	Both	4/4	4/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA
KMDQLAKELTAE	Both	4/4	3/4	12	P10645	344-355	CHGA		Chromogranin A	Y	NA	NA
MDQLAKEL	Both	1/4	4/4	8	P10645	345-352	CHGA		Chromogranin A	Y	NA	NA
MDQLAKELT	Basal	1/4	0/4	9	P10645	345-353	CHGA		Chromogranin A	Y	NA	NA
MDQLAKELTAE	Basal	1/4	0/4	11	P10645	345-355	CHGA		Chromogranin A	Y	NA	NA
DQLAKELT	Basal	1/4	0/4	8	P10645	346-353	CHGA		Chromogranin A	Y	NA	NA
DQLAKELTAE	Both	1/4	2/4	10	P10645	346-355	CHGA		Chromogranin A	Y	NA	NA
LAKELTAE	Both	2/4	2/4	8	P10645	348-355	CHGA		Chromogranin A	Y	NA	NA
LEGQEEEEDNRD	Both	3/4	3/4	12	P10645	358-369	CHGA		Chromogranin A	Y	NA	NA
LEGQEEEEDNRDS	Basal	3/4	0/4	13	P10645	358-370	CHGA		Chromogranin A	Y	NA	NA
LEGQEEEEDNRDSS	Basal	2/4	0/4	14	P10645	358-371	CHGA		Chromogranin A	Y	NA	NA
RAYGFRGP	Both	4/4	4/4	8	P10645	379-386	CHGA		Chromogranin A	Y	NA	NA
RAYGFRGPG	Basal	2/4	0/4	9	P10645	379-387	CHGA		Chromogranin A	Y	NA	NA
RAYGFRGPGP	Both	4/4	4/4	10	P10645	379-388	CHGA		Chromogranin A	Y	NA	NA
RAYGFRGPGPQ	Both	4/4	4/4	11	P10645	379-389	CHGA		Chromogranin A	Y	NA	NA
RAYGFRGPGPQL	Both	4/4	4/4	12	P10645	379-390	CHGA		Chromogranin A	Y	C03:04	0.8606
RAYGFRGPGPQLR	Both	4/4	4/4	13	P10645	379-391	CHGA		Chromogranin A	Y	A03:01	1.7153
AYGFRGPGPQ	Basal	2/4	0/4	10	P10645	380-389	CHGA		Chromogranin A	Y	NA	NA
AYGFRGPGPQL	Both	3/4	4/4	11	P10645	380-390	CHGA		Chromogranin A	Y	NA	NA
YGFRGPGPQL	Both	4/4	4/4	10	P10645	381-390	CHGA		Chromogranin A	Y	C03:04	0.4651
GFRGPGPQ	Both	4/4	3/4	8	P10645	382-389	CHGA		Chromogranin A	Y	NA	NA
GFRGPGPQL	Both	4/4	4/4	9	P10645	382-390	CHGA		Chromogranin A	Y	C07:01	1.3265
GFRGPGPQLR	Both	4/4	4/4	10	P10645	382-391	CHGA		Chromogranin A	Y	NA	NA
FRGPGPQL	Both	4/4	4/4	8	P10645	383-390	CHGA		Chromogranin A	Y	C07:01	0.1921
FRGPGPQLR	Both	3/4	4/4	9	P10645	383-391	CHGA		Chromogranin A	Y	C07:01	0.5683
RGPGPQLR	Basal	1/4	0/4	8	P10645	384-391	CHGA		Chromogranin A	Y	NA	NA
GWRPSSREDS	Basal	1/4	0/4	10	P10645	393-402	CHGA		Chromogranin A	Y	NA	NA
GWRPSSREDSLEA	Basal	1/4	0/4	13	P10645	393-405	CHGA		Chromogranin A	Y	NA	NA
SSREDSLE	Basal	1/4	0/4	8	P10645	397-404	CHGA		Chromogranin A	Y	NA	NA
SSREDSLEA	Both	4/4	1/4	9	P10645	397-405	CHGA		Chromogranin A	Y	NA	NA

Table S2, continued

SSREDSLEAGL	Basal	1/4	0/4	11	P10645	397-407	CHGA		Chromogranin A	Y	B40:01	0.8462
SSREDSLEAGLP	Both	1/4	2/4	12	P10645	397-408	CHGA		Chromogranin A	Y	NA	NA
SSREDSLEAGLPL	Basal	1/4	0/4	13	P10645	397-409	CHGA		Chromogranin A	Y	NA	NA
SSREDSLEAGLPLQ	Both	1/4	2/4	14	P10645	397-410	CHGA		Chromogranin A	Y	NA	NA
SREDSLEA	Both	4/4	4/4	8	P10645	398-405	CHGA		Chromogranin A	Y	NA	NA
SREDSLEAGLP	Both	1/4	1/4	11	P10645	398-408	CHGA		Chromogranin A	Y	NA	NA
SREDSLEAGLPLQ	Both	1/4	1/4	13	P10645	398-410	CHGA		Chromogranin A	Y	NA	NA
SREDSLEAGLPLQV	Both	3/4	4/4	14	P10645	398-411	CHGA		Chromogranin A	Y	NA	NA
REDSLEAGL	IFN-α	0/4	4/4	9	P10645	399-407	CHGA		Chromogranin A	Y	B40:01	0.0618
REDSLEAGLP	IFN-α	0/4	1/4	10	P10645	399-408	CHGA		Chromogranin A	Y	NA	NA
REDSLEAGLPL	IFN-α	0/4	4/4	11	P10645	399-409	CHGA		Chromogranin A	Y	B40:01	0.1959
REDSLEAGLPLQ	Both	1/4	2/4	12	P10645	399-410	CHGA		Chromogranin A	Y	NA	NA
REDSLEAGLPLQV	Both	4/4	4/4	13	P10645	399-411	CHGA		Chromogranin A	Y	B49:01	1.4346
REDSLEAGLPLQVR	Basal	1/4	0/4	14	P10645	399-412	CHGA		Chromogranin A	Y	NA	NA
EDSLEAGLP	Basal	1/4	0/4	9	P10645	400-408	CHGA		Chromogranin A	Y	NA	NA
EDSLEAGLPLQ	Basal	1/4	0/4	11	P10645	400-410	CHGA		Chromogranin A	Y	NA	NA
EDSLEAGLPLQVR	Basal	1/4	0/4	13	P10645	400-412	CHGA		Chromogranin A	Y	NA	NA
SLEAGLPLQVR	Basal	1/4	0/4	11	P10645	402-412	CHGA		Chromogranin A	Y	NA	NA
LEAGLPLQV	Both	2/4	3/4	9	P10645	403-411	CHGA		Chromogranin A	Y	B49:01	0.0173
LEAGLPLQVR	Basal	1/4	0/4	10	P10645	403-412	CHGA		Chromogranin A	Y	NA	NA
EAGLPLQV	Basal	1/4	0/4	8	P10645;P19419	404-411	CHGA;ELK1		Chromogranin A;ELK1, ETS transcription factor	Y	B49:01	1.1357
GYPEEKKEEEGSAN	Both	3/4	3/4	14	P10645	413-426	CHGA		Chromogranin A	Y	NA	NA
RRPEDQELESLS	Basal	1/4	0/4	12	P10645	427-438	CHGA		Chromogranin A	Y	NA	NA
SAIEAELEKV	Both	1/4	1/4	10	P10645	438-447	CHGA		Chromogranin A	Y	NA	NA
AIEAELEKV	IFN-α	0/4	1/4	9	P10645	439-447	CHGA		Chromogranin A	Y	A02:01	1.5618
AIEAELEKVAHQ	Basal	1/4	0/4	12	P10645	439-450	CHGA		Chromogranin A	Y	NA	NA
IEAELEKV	IFN-α	0/4	4/4	8	P10645	440-447	CHGA		Chromogranin A	Y	B49:01	0.2445
IEAELEKVA	IFN-α	0/4	2/4	9	P10645	440-448	CHGA		Chromogranin A	Y	B49:01	0.3508
AELEKVAHQL	IFN-α	0/4	3/4	10	P10645	442-451	CHGA		Chromogranin A	Y	B40:01	0.0394
KVAHQLQAL	Both	1/4	1/4	9	P10645	446-454	CHGA		Chromogranin A	Y	C03:04	0.487
KVAHQLQALR	Basal	2/4	0/4	10	P10645	446-455	CHGA		Chromogranin A	Y	A03:01	1.0954
GLEPGKGRHH	Both	4/4	4/4	10	P05060	410-419	CHGB	SCG1,SgI	Chromogranin B	Y	NA	NA
EPGKGRHH	Basal	2/4	0/4	8	P05060	412-419	CHGB	SCG1,SgI	Chromogranin B	Y	NA	NA
RVAQLDQLLH	Both	4/4	2/4	10	P05060	603-612	CHGB	SCG1,SgI	Chromogranin B	Y	A03:01	0.8134
RVAQLDQLLHY	Both	4/4	4/4	11	P05060	603-613	CHGB	SCG1,SgI	Chromogranin B	Y	A03:01	0.6234
AQLDQLLHY	Both	4/4	3/4	9	P05060	605-613	CHGB	SCG1,SgI	Chromogranin B	Y	C07:01/A03:01	0.309/0.715
KIAEKFSQR	Basal	2/4	0/4	9	P05060	668-676	CHGB	SCG1,SgI	Chromogranin B	Y	A03:01	0.098
KPKAAPKPKFSWSK	Basal	1/4	0/4	14	Q12860	437-450	CNTN1	F3,GP135	Contactin 1	N	NA	NA
KVLYRPDGQHDGK	Basal	2/4	0/4	13	Q12860	941-953	CNTN1	F3,GP135	Contactin 1	N	NA	NA
VLYRPDGQHDGK	Both	4/4	4/4	12	Q12860	942-953	CNTN1	F3,GP135	Contactin 1	N	A03:01	1.2858
EEYGKLQKI	IFN-α	0/4	2/4	9	P61201	163-171	COPS2	ALIEN,CSN2,TRIF	COP9 signalosome subunit 2	N	B49:01	0.0312
ALYEQSLHIK	Basal	1/4	0/4	10	P61201	216-225	COPS2	ALIEN,CSN2,TRIF	COP9 signalosome subunit 2	N	A03:01	0.0101
SAIPHPLIM	IFN-α	0/4	4/4	9	P61201	226-234	COPS2	ALIEN,CSN2,TRIF	COP9 signalosome subunit 2	N	C03:04	0.0019
TEFEKILKT	IFN-α	0/4	2/4	9	P61201	324-332	COPS2	ALIEN,CSN2,TRIF	COP9 signalosome subunit 2	N	B49:01	0.1132
SLNQAVVSK	Both	3/4	4/4	9	P61201	433-441	COPS2	ALIEN,CSN2,TRIF	COP9 signalosome subunit 2	N	A03:01	0.0091
LINKSGELL	IFN-α	0/4	1/4	9	Q9UNS2	27-35	COPS3	CSN3,SGN3	COP9 signalosome subunit 3	N	C03:04	1.4554
SLYKKNIQR	Both	4/4	1/4	9	Q9UNS2	301-309	COPS3	CSN3,SGN3	COP9 signalosome subunit 3	N	A03:01	0.0334
TVNPQFVQK	Both	3/4	4/4	9	Q9UNS2	398-406	COPS3	CSN3,SGN3	COP9 signalosome subunit 3	N	A03:01	0.0111
TAISRIYTV	IFN-α	0/4	1/4	9	P16870	71-79	CPE		Carboxypeptidase E	Y	C03:04/A02:01	0.2424/0.4707
RSFEGREL	Both	2/4	2/4	8	P16870	81-88	CPE		Carboxypeptidase E	Y	C03:04	0.4425
RSFEGRELL	IFN-α	0/4	1/4	9	P16870	81-89	CPE		Carboxypeptidase E	Y	E01:01	0.0736
FEGRELLVI	IFN-α	0/4	4/4	9	P16870	83-91	CPE		Carboxypeptidase E	Y	B49:01/B40:01	0.2671/0.5285

Table S2, continued

RNFPDLDRI	IFN-α	0/4	1/4	9	P16870	189-197	CPE		Carboxypeptidase E	Y	B49:01	1.8692
KEGGPNNHL	IFN-α	0/4	3/4	9	P16870	203-211	CPE		Carboxypeptidase E	Y	B40:01	0.0911
KEGGPNNHLL	IFN-α	0/4	4/4	10	P16870	203-212	CPE		Carboxypeptidase E	Y	B40:01	0.278
WEDNKNSLI	IFN-α	0/4	2/4	9	P16870	358-366	CPE		Carboxypeptidase E	Y	B49:01/B40:01	0.262/0.4403
RLLIPGNYK	Both	3/4	4/4	9	P16870	412-420	CPE		Carboxypeptidase E	Y	A03:01	0.0132
RLLIPGNYKL	Both	1/4	1/4	10	P16870	412-421	CPE		Carboxypeptidase E	Y	A02:01/A03:01	0.2705/0.6087
ASAPGYLAITK	Both	4/4	4/4	11	P16870	423-433	CPE		Carboxypeptidase E	Y	A03:01	0.2238
KMMSETLNF	Both	4/4	4/4	9	P16870	468-476	CPE		Carboxypeptidase E	Y	E01:01	0.1764
KEEERQEAL	IFN-α	0/4	3/4	9	Q6PUV4;O14810	33-41	CPLX2;CPLX1	CPX-2,DKFZp547	Complexin 2;Complexin 1	N	B40:01	0.0351
AEAVFRAI	Basal	1/4	0/4	8	Q14194;Q14195	233-240	CRMP1;DPYSI	DPYSL1,DRP-1,C	Collapsin response mediator protein 1;Dihydrop	N	B49:01	0.1211
KMDENQFVA	IFN-α	0/4	1/4	9	Q14194;Q16555	374-382	CRMP1;DPYSI	DPYSL1,DRP-1,C	Collapsin response mediator protein 1;Dihydrop	N	A02:01	0.3122
IVAPPGGRSNITS	Basal	1/4	0/4	13	Q14194;Q14195	558-570	CRMP1;DPYSI	DPYSL1,DRP-1,C	Collapsin response mediator protein 1;Dihydrop	N	NA	NA
RTAGAIRGR	Basal	4/4	0/4	9	P35221;P26232	540-548	CTNNA1;CTN	CAP102,CAP-R,C	Catenin alpha 1;Catenin alpha 2	N	A03:01	0.3446
ASYVASTKYQK	Basal	1/4	0/4	11	P35221;P26232	835-845	CTNNA1;CTN	CAP102,CAP-R,C	Catenin alpha 1;Catenin alpha 2	N	A03:01	0.2798
KMKAPEKKPLVK	Basal	4/4	0/4	12	P35221;P26232	860-871	CTNNA1;CTN	CAP102,CAP-R,C	Catenin alpha 1;Catenin alpha 2	N	A03:01	0.272
KMVDLRGAKV	IFN-α	0/4	3/4	10	Q9UI36	189-198	DACH1	DACH	Dachshund family transcription factor 1	N	A02:01	1.4357
HLVGGLHTV	IFN-α	0/4	2/4	9	Q9UI36;Q96NX9	221-229	DACH1;DACH	DACH,	Dachshund family transcription factor 1;Dachsh	N	A02:01	0.01
AIAEAMKVK	Both	2/4	3/4	9	Q9UI36	337-345	DACH1	DACH	Dachshund family transcription factor 1	N	A03:01	0.1576
AEAMKVKKI	IFN-α	0/4	3/4	9	Q9UI36	339-347	DACH1	DACH	Dachshund family transcription factor 1	N	B49:01	0.0689
GLSPNVLPGPK	Basal	1/4	0/4	11	Q9UI36	545-555	DACH1	DACH	Dachshund family transcription factor 1	N	A03:01	0.363
RERELRETL	IFN-α	0/4	4/4	9	Q9UI36	656-664	DACH1	DACH	Dachshund family transcription factor 1	N	B40:01	0.0416
RLDLGEDYPSGK	Basal	1/4	0/4	12	O43143	6-17	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	A03:01	1.1234
RLDLGEDYPSGKK	Basal	1/4	0/4	13	O43143	6-18	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	NA	NA
LDLGEDYPSGK	Basal	1/4	0/4	11	O43143	7-17	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	NA	NA
LDLGEDYPSGKK	Basal	1/4	0/4	12	O43143	7-18	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	NA	NA
AGHAGHTSL	Basal	1/4	0/4	9	O43143	103-111	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	C03:04	0.5691
ILMGVLKEV	Both	2/4	2/4	9	O43143	271-279	DHX15	DBP1,DDX15,HR	DEAH;box helicase 15	N	A02:01	0.0154
LEAAIRTV	IFN-α	0/4	1/4	8	O43143	332-339	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	B49:01	0.1425
RIFEPPPPK	IFN-α	0/4	2/4	9	O43143	399-407	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	A03:01	0.0022
VVFVIDPGFAK	IFN-α	0/4	3/4	11	O43143	433-443	DHX15	DBP1,DDX15,HR	DEAH-box helicase 15	N	A03:01	0.1159
REEGPAETV	IFN-α	0/4	2/4	9	Q9BTC0	193-201	DIDO1	BYE1,C20orf158	Death inducer-obliterator 1	N	B49:01/B40:01	0.044/0.0905
SVHKRPAPEK	IFN-α	0/4	1/4	10	Q9BTC0	465-474	DIDO1	BYE1,C20orf158	Death inducer-obliterator 1	N	A03:01	0.1557
KTAPPGSAV	IFN-α	0/4	2/4	9	Q9BTC0	549-557	DIDO1	BYE1,C20orf158	Death inducer-obliterator 1	N	C03:04	0.268
KTAPPGSAVGK	Both	4/4	4/4	11	Q9BTC0	549-559	DIDO1	BYE1,C20orf158	Death inducer;obliterator 1	N	A03:01	0.0101
AAATPAIKK	Both	1/4	4/4	9	Q9BTC0	577-585	DIDO1	BYE1,C20orf158	Death inducer;obliterator 1	N	A03:01	0.0826
VMPEAVPEV	IFN-α	0/4	1/4	9	Q9BTC0	902-910	DIDO1	BYE1,C20orf158	Death inducer-obliterator 1	N	A02:01	0.0088
SLYSYFSSR	IFN-α	0/4	2/4	9	Q9BTC0	1135-1143	DIDO1	BYE1,C20orf158	Death inducer-obliterator 1	N	A03:01	0.0787
TEALLRVL	IFN-α	0/4	1/4	8	P80370	4-11	DLK1	Delta1,FA1,pG2,I	Delta like non-canonical Notch ligand 1	N	B40:01/B49:01	0.4387/0.6949
TEALLRVLL	IFN-α	0/4	4/4	9	P80370	4-12	DLK1	Delta1,FA1,pG2,I	Delta like non-canonical Notch ligand 1	N	B40:01	0.0495
AAIDHQIEV	IFN-α	0/4	2/4	9	Q9Y485	473-481	DMXL1		Dmx like 1	N	C03:04	0.0602
REDGSHILTV	IFN-α	0/4	4/4	10	Q9Y485	1147-1156	DMXL1		Dmx like 1	N	B49:01/B40:01	0.0829/0.2229
SSSGLHPPK	Both	2/4	4/4	9	Q9Y485	1271-1279	DMXL1		Dmx like 1	N	A03:01	0.2813
GLFEAAHVL	IFN-α	0/4	2/4	9	Q9Y485;Q8TDJ6	1310-1318	DMXL1;DMXL	KIAA0856,RC3	Dmx like 1;Dmx like 2	N	A02:01	0.0336
AAYKSILRK	IFN-α	0/4	4/4	9	Q9Y485	1739-1747	DMXL1		Dmx like 1	N	A03:01	0.0107
HERFPFELV	IFN-α	0/4	4/4	9	P50570;Q05193	367-375	DNM2;DNM1	CMT2M,CMTDI1	Dynamin 2;Dynamin 1	N	B49:01	0.0893
NHEDFIGF	Basal	1/4	0/4	8	P50570;Q05193;C	489-496	DNM2;DNM1	CMT2M,CMTDI1	Dynamin 2;Dynamin 1;Dynamin 3	N	NA	NA
ITSDRLLIK	Both	3/4	4/4	9	Q16555;Q14195;(	12-20	DPYSL2;DPYSI	CRMP2,CRMP4,I	Collapsin response mediator protein 1;Dihydrop	N	A03:01	0.107
MEDGLIKQI	IFN-α	0/4	2/4	9	Q16555	37-45	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	B49:01	0.0126
IAEEQQRIL	Both	1/4	3/4	9	Q16555	205-213	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	C03:04	0.0576
AEAVNRAITI	IFN-α	0/4	3/4	10	Q16555	233-242	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	B49:01/B40:01	0.25/0.647
KVFNLYPRKG	IFN-α	0/4	2/4	10	Q16555	390-399	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	A03:01	1.8034
VIWDPDSVKTISA	Both	4/4	4/4	13	Q16555	410-422	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA

Table S2, continued

KTHNSSLEY	Both	3/4	4/4	9	Q16555	423-431	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	A03:01/C07:01	0.1401/0.2128
RLAELRGVPR	Both	3/4	4/4	10	Q16555	487-496	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	A03:01	0.604
SLSGAQIDDNIPR	Both	3/4	4/4	13	Q16555	540-552	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
SLSGAQIDDNIPRR	Basal	1/4	0/4	14	Q16555	540-553	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
LSGAQIDDNIPR	Basal	1/4	0/4	12	Q16555	541-552	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
AQIDDNIPRRTTQ	Basal	1/4	0/4	13	Q16555	544-556	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
IDDNIPRRT	Basal	1/4	0/4	9	Q16555	546-554	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
IDDNIPRRTTQ	Basal	1/4	0/4	11	Q16555	546-556	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
RIVAPPGGR	Both	4/4	4/4	9	Q16555;Q14195;(	557-565	DPYSL2;DPYSI	CRMP2,CRMP4,[	Collapsin response mediator protein 1;Dihydro	N	A03:01	0.3309
IVAPPGGRANITS	Basal	1/4	0/4	13	Q16555	558-570	DPYSL2	CRMP2,DHPRP2,	Dihydropyrimidinase like 2	N	NA	NA
YTSKLVKTL	IFN-α	0/4	2/4	9	Q7RTS9	210-218	DYM	DMC,FLJ20071,S	Dymeclin	N	C03:04	0.0832
GSYVTHETK	Both	2/4	4/4	9	Q96FJ2;P63167	63-71	DYNLL1;DYNL	DLC1,DLC8,DNC	Dynein light chain LC8;type 1;Dynein light chain	N	A03:01	0.4961
SELITPHAI	IFN-α	0/4	2/4	9	Q9UH73;Q9H4W6	302-310	EBF1;EBF3;EB	EBF,OLF1,COE3,I	Early B-cell factor 1;Early B-cell factor 3;Early E	N	B49:01	0.0094
KITGQSLGY	Basal	2/4	0/4	9	Q14576;P26378;C	74-82	ELAVL3;ELAVI	DKFZp547J036,H	ELAV like RNA binding protein 3;ELAV like RNA I	N	A03:01	0.361
ILVDQVTGV	Basal	1/4	0/4	9	Q14576;P26378	156-164	ELAVL3;ELAVI	DKFZp547J036,H	ELAV like RNA binding protein 3;ELAV like RNA I	N	A02:01	0.0083
NLLEKDYFGL	IFN-α	0/4	1/4	10	Q9Y2J2;Q9H4G0;(	144-153	EPB41L3;EPB4	DAL1,KIAA0987,4	Erythrocyte membrane protein band 4.1 like 3;I	N	A02:01	0.7314
SEDITRYYL	IFN-α	0/4	3/4	9	Q9Y2J2	197-205	EPB41L3	DAL1,KIAA0987	Erythrocyte membrane protein band 4.1 like 3	N	B40:01/B49:01	0.0824/0.2149
KELEDKVIEL	IFN-α	0/4	1/4	10	Q9Y2J2	261-270	EPB41L3	DAL1,KIAA0987	Erythrocyte membrane protein band 4.1 like 3	N	B40:01	0.0464
SMYGVDLHHA	Both	1/4	3/4	10	O43491;Q9H4G0;(	402-411	EPB41L1;EPB4	DAL1,KIAA0338,I	Erythrocyte membrane protein band 4;1 like 3;I	N	A02:01	0.3101
AAIHISETL	IFN-α	0/4	4/4	9	Q9Y2J2	932-940	EPB41L3	DAL1,KIAA0987	Erythrocyte membrane protein band 4.1 like 3	N	C03:04	0.0351
GISETRIEK	Basal	3/4	0/4	9	O43491;P11171;C	948-956	EPB41;EPB41I	EL1,4.1-G,4.1B,[	Erythrocyte membrane protein band 4.1;Erythr	N	A03:01	0.1604
GSVSPGGVK	Basal	1/4	0/4	9	Q9Y2J2	959-967	EPB41L3	DAL1,KIAA0987	Erythrocyte membrane protein band 4.1 like 3	N	A03:01	0.7498
KIKPRHDSGIK	Basal	1/4	0/4	11	Q9NZ56	1705-1715	FMN2		Formin 2	N	A03:01	0.9222
GEYVIKSLF	IFN-α	0/4	1/4	9	O94915	13-21	FRYL	AF4p12,DKFZp68	FRY like transcription coactivator	N	B40:01/B49:01	0.2205/0.2711
SVYYPQVRK	Both	1/4	4/4	9	O94915	505-513	FRYL	AF4p12,DKFZp68	FRY like transcription coactivator	N	A03:01	0.0032
GLDNETHFL	IFN-α	0/4	1/4	9	O94915	1003-1011	FRYL	AF4p12,DKFZp68	FRY like transcription coactivator	N	A02:01	0.0413
KMFPEVKEK	Both	4/4	4/4	9	Q05329	257-265	GAD2	GAD65	Glutamate decarboxylase 2	Y	A03:01	0.0069
KVAPVIKAR	IFN-α	0/4	4/4	9	Q05329	528-536	GAD2	GAD65	Glutamate decarboxylase 2	Y	A03:01	0.1489
VSLDRYLAI	IFN-α	0/4	1/4	9	O43603	119-127	GALR2	GALNR2	Galanin receptor 2	N	C03:04	0.6528
SEDVMRITV	IFN-α	0/4	3/4	9	P35557	398-406	GCK	HK4,MODY2	Glucokinase	N	B49:01	0.0375
GSSFLSPEH	Basal	1/4	0/4	9	Q9UBU3	24-32	GHRL	ghrelin,MTLRP,o	Ghrelin and obestatin prepropeptide	N	NA	NA
KMFDVGGQR	Basal	3/4	0/4	9	P04899;P08754;P(	198-206	GNAI2;GNAI3	GIP,GNAI2B,87U	G protein subunit alpha i2;G protein subunit alp	N	A03:01	0.1598
GESGKSTIV	Both	4/4	4/4	9	Q5JWF2;P63092;F	692-700	GNAS;GNAI1;	87U6,GDCA,GIP,	G protein subunit alpha L;G protein subunit alpl	N	B49:01/B40:01	0.1615/0.308
LAEKVLAGK	IFN-α	0/4	1/4	9	Q5JWF2;P63092;F	940-948	GNAS;GNAS;C	GNAS1,GNASXL,(	GNAS complex locus;GNAS complex locus;G pro	N	NA	NA
AEKVLAGK	IFN-α	0/4	1/4	8	Q5JWF2;P63092;F	941-948	GNAS;GNAS;C	GNAS1,GNASXL,(	GNAS complex locus;GNAS complex locus;G pro	N	NA	NA
GELERALSAV	IFN-α	0/4	1/4	10	Q08379;I6L899;H:	266-275	GOLGA2;GOL(	GM130,golgin-95	Golgin A2;Golgin A8 family member R;Golgin A8	N	B49:01/B40:01	0.25/0.7259
SVSQINHKL	IFN-α	0/4	2/4	9	Q9HD26	86-94	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	C03:04	0.2668
TETQAEKVVL	IFN-α	0/4	4/4	10	Q9HD26	106-115	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	B40:01	0.1605
KEVHDQLLQL	IFN-α	0/4	2/4	10	Q9HD26	117-126	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	B40:01/B49:01	0.1055/0.2895
HEGLGISI	IFN-α	0/4	4/4	8	Q9HD26	296-303	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	B49:01/B40:01	0.3244/0.8213
KEHGVPILI	IFN-α	0/4	4/4	9	Q9HD26	307-315	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	B49:01	0.0083
GEIKVLQGF	IFN-α	0/4	4/4	9	Q9HD26	413-421	GOPC	CAL,dJ94G16.2,F	Golgi associated PDZ and coiled-coil motif cont	N	B40:01/B49:01	0.2171/0.2757
KLDDLHTLYH	Both	1/4	2/4	10	Q9HD26	449-458	GOPC	CAL,FIG,GOPC1,I	Golgi associated PDZ and coiled;coil motif cont	N	A03:01	0.4068
SEYPIVDGKL	Both	3/4	4/4	10	P54868;Q01581	248-257	HMGCS2;HM(	HMGCS	hydroxy;methylglutaryl;CoA synthase 1;hydroxy	N	B40:01/B49:01	0.0853/0.2371
SEYPIVDGKLSI	Both	2/4	4/4	12	P54868;Q01581	248-259	HMGCS2;HM(	HMGCS	hydroxy;methylglutaryl;CoA synthase 1;hydroxy	N	B49:01	0.1032
RVFIGNLNTLVVK	Both	4/4	4/4	13	P0DMR1;B7ZW38	17-29	HNRNPCL4;H	HNRNPCP5,HNR	Heterogeneous nuclear ribonucleoprotein C;like	N	A03:01	0.3462
IGNLNTLVVK	IFN-α	0/4	1/4	10	P0DMR1;P07910;I	20-29	HNRNPCL4;H	hnRNPC,HNRPC,	Heterogeneous nuclear ribonucleoprotein C (C:	N	NA	NA
SVHKGFAFV	Both	1/4	3/4	9	P0DMR1;B7ZW38	47-55	HNRNPCL4;H	HNRNPCP5,HNR	Heterogeneous nuclear ribonucleoprotein C;like	N	A02:01	0.8047
RVPPPPPIA	Both	3/4	2/4	9	P0DMR1;B7ZW38	129-137	HNRNPCL4;H	HNRNPCP5,HNR	Heterogeneous nuclear ribonucleoprotein C;like	N	A02:01	1.3086
SLLENLEKI	Both	2/4	4/4	9	P0DMR1;B7ZW38	196-204	HNRNPCL4;H	HNRNPCP5,HNR	Heterogeneous nuclear ribonucleoprotein C;like	N	A02:01	0.0209
YESLTDPSKL	IFN-α	0/4	1/4	10	P08238;P07900;Q	56-65	HSP90AB1;HSP90AA1;HSP90AI		Heat shock protein 90 alpha family class B mem	N	B40:01	0.137
KEFDGKSL	IFN-α	0/4	2/4	8	P08238;Q58FF7;Q	526-533	HSP90AB1;HSP90AB3P;HSP90A		Heat shock protein 90 alpha family class B mem	N	B40:01/B49:01	0.1249/0.2217

Table S2, continued

VTKEGLEL	IFN-α	0/4	2/4	8	P08238;P07900;Q	536-543	HSP90AB1;HSP	90AA1;HSP90AE	Heat shock protein 90 alpha family class B mem	N	C03:04	0.6051
KLSLVAAML	Both	3/4	3/4	9	P11021	2-10	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A02:01	0.8236
TVFDAKRLIGR	Both	4/4	4/4	11	P11021;P17066;P(	91-101	HSPA5;HSPA6	BiP,GRP78,HSC7(	Heat shock protein family A (Hsp70) member 8;	N	A03:01	1.5997
NDPSVQQDIKFLP	Basal	1/4	0/4	13	P11021	104-116	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	NA	NA
AEAYLGKKV	Both	2/4	4/4	9	P11021	157-165	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	B49:01	0.0216
EAYLGKKV	IFN-α	0/4	1/4	8	P11021	158-165	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	NA	NA
RIINEPTAA	Basal	1/4	0/4	9	P11021;P11142;P(	197-205	HSPA5;HSPA8	HSC70,HSC71,HS	Heat shock protein family A (Hsp70) member 8;	N	A02:01	1.598
RIINEPTAAAIAY	Basal	1/4	0/4	13	P11021;P11142;P(	197-209	HSPA5;HSPA8	HSC70,HSC71,HS	Heat shock protein family A (Hsp70) member 8;	N	NA	NA
RIINEPTAAAIAYG	Both	1/4	1/4	14	P11021;P17066;P(	197-210	HSPA5;HSPA6	BiP,GRP78,HSC7(	Heat shock protein family A (Hsp70) member 8;	N	NA	NA
EGEKNILV	IFN-α	0/4	1/4	8	P11021	215-222	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	NA	NA
RVMEHFIKL	IFN-α	0/4	1/4	9	P11021	261-269	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A02:01	0.0432
KLIPRNTVVPTK	IFN-α	0/4	2/4	12	P11021	435-446	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A03:01	0.1272
KVYEGERPLTK	IFN-α	0/4	3/4	11	P11021	464-474	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A03:01	0.0031
VEVLKREPL	IFN-α	0/4	2/4	9	P10997	76-84	IAPP	AMYLIN,DAP,IAF	Islet amyloid polypeptide	Y	B40:01	0.2693
ALWGPDPA	Both	2/4	1/4	8	P01308;F8WCM5	15-22	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA
ALWGPDPAA	Basal	2/4	0/4	9	P01308;F8WCM5	15-23	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	A02:01	0.2955
ALWGPDPAAA	Both	4/4	4/4	10	P01308;F8WCM5	15-24	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.3268
AAFVNQHL	IFN-α	0/4	1/4	8	P01308;F8WCM5	23-30	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	C03:04	0.5183
HLCGSHLVEA	IFN-α	0/4	1/4	10	P01308;F8WCM5	29-38	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA
GSHLVEALYL	IFN-α	0/4	1/4	10	P01308;F8WCM5	32-41	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA
HLVEALYL	Both	4/4	4/4	8	P01308;F8WCM5	34-41	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA
HLVEALYLV	Both	4/4	4/4	9	P01308;F8WCM5	34-42	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.0364
GERGFFYTP	Both	3/4	4/4	9	P01308;F8WCM5	44-52	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	B49:01/B40:01	0.4123/1.2003
GERGFFYTPK	Both	1/4	1/4	10	P01308;F8WCM5	44-53	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA
GERGFFYTPKT	Both	1/4	2/4	11	P01308;F8WCM5	44-54	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA
ERGFFYTPK	Basal	1/4	0/4	9	P01308;F8WCM5	45-53	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA
GFFYTPKT	Basal	1/4	0/4	8	P01308;F8WCM5	47-54	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA
REAEDLQV	Both	1/4	4/4	8	P01308	56-63	INS	IDDM1,IDDM2	Insulin	Y	B49:01	0.1817
REAEDLQVG	IFN-α	0/4	1/4	9	P01308	56-64	INS	IDDM1,IDDM2	Insulin	Y	B49:01	1.1569
REAEDLQVGQV	Both	1/4	4/4	11	P01308	56-66	INS	IDDM1,IDDM2	Insulin	Y	B49:01/B40:01	0.1983/0.5284
REAEDLQVGQVE	Basal	1/4	0/4	12	P01308	56-67	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
REAEDLQVGQVEL	IFN-α	0/4	4/4	13	P01308	56-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.2255
EAEDLQVGQVELGG	Basal	1/4	0/4	14	P01308	57-70	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
AEDLQVGQ	Basal	1/4	0/4	8	P01308	58-65	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
AEDLQVGQV	IFN-α	0/4	3/4	9	P01308	58-66	INS	IDDM1,IDDM2	Insulin	Y	B49:01	0.0989
AEDLQVGQVEL	Both	2/4	4/4	11	P01308	58-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.1346
LQVGQVEL	IFN-α	0/4	1/4	8	P01308	61-68	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
VELGGGPGAGS	Basal	1/4	0/4	11	P01308	66-76	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
VELGGGPGAGSL	Both	2/4	4/4	12	P01308	66-77	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.327
VELGGGPGAGSLQP	Both	2/4	1/4	14	P01308	66-79	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
GGGPGAGSL	Both	3/4	4/4	9	P01308	69-77	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
AGSLQPLALEGSLQ	Basal	2/4	0/4	14	P01308	74-87	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
GSLQPLALEGSLQ	IFN-α	0/4	2/4	13	P01308	75-87	INS	IDDM1,IDDM2	Insulin	Y	NA	NA
TVLNEKQLHTL	IFN-α	0/4	2/4	11	P61371;Q96A47	186-196	ISL1;ISL2	Isl-1,ISLET1,FLJ1	ISL LIM homeobox 1;ISL LIM homeobox 2	N	NA	NA
YAANPRPDAL	IFN-α	0/4	1/4	10	P61371;Q96A47	200-209	ISL1;ISL2	Isl-1,ISLET1,FLJ1	ISL LIM homeobox 1;ISL LIM homeobox 2	N	C03:04	0.0863
AANPRPDAL	Both	2/4	4/4	9	P61371;Q96A47	201-209	ISL1;ISL2	FLJ10160,ISLET1	ISL LIM homeobox 1;ISL LIM homeobox 2	N	C03:04	0.0326
KLTDIQIEL	Both	3/4	4/4	9	Q96AA8;Q5VZ66	29-37	JAKMIP2;JAKM	C10orf14,C10orf	Janus kinase and microtubule interacting protei	N	A02:01	0.0045
RVLPLLLAY	IFN-α	0/4	3/4	9	Q96T55	13-21	KCNK16	K2p16.1,TALK-1,	Potassium two pore domain channel subfamily	N	A03:01	0.0834
SEYHYEYTA	IFN-α	0/4	1/4	9	Q6UXG2	52-60	KIAA1324	EIG121,maba1	KIAA1324	N	B49:01	0.1246
VAVPHTPGL	Both	1/4	4/4	9	Q6UXG2	70-78	KIAA1324	EIG121,maba1	KIAA1324	N	C03:04	0.0051
WEFHSVEL	IFN-α	0/4	1/4	8	Q6UXG2	228-235	KIAA1324	EIG121,maba1	KIAA1324	N	B40:01	0.1458
GESGFSKSI	IFN-α	0/4	3/4	9	Q6UXG2	726-734	KIAA1324	EIG121,maba1	KIAA1324	N	B49:01	0.0237

Table S2, continued

SLADRLIGV	IFN-α	0/4	3/4	9	Q6UXG2	761-769	KIAA1324	EIG121,maba1	KIAA1324	N	A02:01	0.0071
AELFHLESL	IFN-α	0/4	1/4	9	Q6UXG2	782-790	KIAA1324	EIG121,maba1	KIAA1324	N	B40:01	0.0183
KTIDFWLKV	IFN-α	0/4	1/4	9	Q6UXG2	903-911	KIAA1324	EIG121,maba1	KIAA1324	N	A02:01/C07:01	0.2555/0.6864
SLFGKIKSFTSK	Both	4/4	4/4	12	Q6UXG2	980-991	KIAA1324	EIG121,maba1	KIAA1324	N	A03:01	0.0767
RVRDLLNPK	Both	2/4	4/4	9	Q12756;O60333	153-161	KIF1A;KIF1B	ATSV,C2orf20,CI	Kinesin family member 1A;Kinesin family memb	N	A03:01	0.0377
RVRDLLNPKNK	IFN-α	0/4	4/4	11	Q12756;O60333	153-163	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	A03:01	0.2463
AVFNIIFTQK	Both	4/4	4/4	10	Q12756	219-228	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.022
TEKVSKISL	IFN-α	0/4	4/4	9	Q12756;O60333	238-246	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	B40:01/B49:01	0.1422/0.2536
RLKEGANINK	Both	4/4	4/4	10	Q12756;O60333	264-273	KIF1A;KIF1B	ATSV,C2orf20,CI	Kinesin family member 1A;Kinesin family memb	N	A03:01	0.0224
KEGANINKSL	IFN-α	0/4	1/4	10	Q12756;O60333	266-275	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	B40:01	0.2154
SLHERILFA	IFN-α	0/4	1/4	9	Q12756	419-427	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A02:01	0.0803
ILFAPGSEEAIER	Basal	1/4	0/4	13	Q12756	424-436	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	NA	NA
KETEKIIAEL	IFN-α	0/4	4/4	10	Q12756	438-447	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B40:01	0.0979
REDGGTLGV	Both	2/4	4/4	9	Q12756;O60333	477-485	KIF1A;KIF1B	ATSV,C2orf20,CI	Kinesin family member 1A;Kinesin family memb	N	B49:01/B40:01	0.1183/0.2492
REDGGTLGVF	IFN-α	0/4	4/4	10	Q12756;O60333	477-486	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	B40:01/B49:01	0.2566/0.622
YIKDGITRV	Both	3/4	4/4	9	Q12756;O60333	510-518	KIF1A;KIF1B	ATSV,C2orf20,CI	Kinesin family member 1A;Kinesin family memb	N	A02:01	0.0732
REREEATYL	IFN-α	0/4	4/4	9	Q12756	651-659	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B40:01/B49:01	0.0357/0.0795
KEANAISV	IFN-α	0/4	2/4	8	Q12756;O60333	734-741	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	B49:01	0.1348
KEANAISVEL	Both	1/4	4/4	10	Q12756;O60333	734-743	KIF1A;KIF1B	ATSV,C2orf20,CI	Kinesin family member 1A;Kinesin family memb	N	B40:01	0.057
ATHYWTLEKL	IFN-α	0/4	4/4	10	Q12756	796-805	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	NA	NA
AEVPSSVIE	IFN-α	0/4	1/4	9	Q12756	820-828	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B40:01	1.7796
LLYPVPLVH	Both	4/4	4/4	9	Q12756	860-868	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.0638
LLYPVPLVHR	Both	2/4	4/4	10	Q12756	860-869	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.1839
GEVKGFLRV	Both	2/4	4/4	9	Q12756	877-885	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B49:01	0.0161
RLGNTFTFR	Both	1/4	3/4	9	Q12756	995-1003	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.2336
IRWKEVREL	IFN-α	0/4	4/4	9	Q12756	1189-1197	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	C07:01	0.0221
RELVVGRI	IFN-α	0/4	4/4	8	Q12756;O60333	1195-1202	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	B49:01	0.2966
RVLDTSVAY	Both	1/4	1/4	9	Q12756	1346-1354	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.1778
RVLDTSVAYV	IFN-α	0/4	3/4	10	Q12756	1346-1355	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A02:01	0.4568
VLDTSVAYV	IFN-α	0/4	1/4	9	Q12756	1347-1355	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A02:01	0.0238
RLLTHTFNR	IFN-α	0/4	3/4	9	Q12756	1457-1465	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.0959
REYTHSHV	IFN-α	0/4	2/4	8	Q12756	1465-1472	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B49:01	0.086
KLSEMSVTL	Both	3/4	4/4	9	Q12756	1480-1488	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A02:01	0.0117
RVSPIVSKK	Both	4/4	4/4	9	Q12756	1571-1579	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	A03:01	0.0037
TEHRGILL	IFN-α	0/4	4/4	8	Q12756	1646-1653	KIF1A	ATSV,C2orf20,SF	Kinesin family member 1A	N	B40:01/B49:01	0.3911/0.4607
LLAGTIRSK	IFN-α	0/4	4/4	9	Q12756;O60333	1671-1679	KIF1A;KIF1B	ATSV,C2orf20,SF	Kinesin family member 1A;Kinesin family memb	N	A03:01	0.0805
RRIEVNVEL	Basal	4/4	0/4	9	P52292	30-38	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	C07:01	0.0075
RIEVNVELRK	Basal	1/4	0/4	10	P52292	31-40	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A03:01	0.4176
IEVNVELR	Basal	1/4	0/4	8	P52292	32-39	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	NA	NA
AVFPSLLTNPK	Both	4/4	3/4	11	P52292	338-348	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A03:01	0.0088
LTNPKTNIQK	IFN-α	0/4	1/4	10	P52292	344-353	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A03:01	0.1083
TNPKTNIQKEATW	Basal	1/4	0/4	13	P52292	345-357	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	NA	NA
GTVEQIVYL	IFN-α	0/4	1/4	9	P52292	408-416	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A02:01	0.2869
AEKLGETEKL	IFN-α	0/4	4/4	10	P52292	451-460	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	B40:01	0.1019
KLGETEKL	Basal	2/4	0/4	8	P52292	453-460	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	NA	NA
TEKLSIMI	IFN-α	0/4	3/4	8	P52292	457-464	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	B49:01	0.4212
KASLSLIEK	Both	2/4	2/4	9	P52292	486-494	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A03:01	0.2172
SLIEKYFSV	IFN-α	0/4	1/4	9	P52292	490-498	KPNA2	IPOA1,QIP2,RCH	Karyopherin subunit alpha 2	N	A02:01	0.0021
KTAKLILERF	Basal	1/4	0/4	10	Q9C0E8	130-139	LNPK	KIAA1715,LNP,LI	Lunapark, ER junction formation factor	N	NA	NA
KLRKSSHLK	Both	2/4	3/4	9	Q9Y250	24-32	LZTS1	FEZ1	Leucine zipper tumor suppressor 1	N	A03:01	0.0215
KLRKSSHLKKLN	Basal	1/4	0/4	12	Q9Y250	24-35	LZTS1	FEZ1	Leucine zipper tumor suppressor 1	N	NA	NA
SEDFFYIKV	IFN-α	0/4	3/4	9	Q9Y250	62-70	LZTS1	FEZ1	Leucine zipper tumor suppressor 1	N	B49:01	0.1046

Table S2, continued

VEDRFSDDQL	IFN-α	0/4	1/4	10	Q9Y5Q3	207-216	MAFB	KRML	MAF bZIP transcription factor B	N	B40:01	0.7501
LENEKTQLI	IFN-α	0/4	4/4	9	Q9Y5Q3	266-274	MAFB	KRML	MAF bZIP transcription factor B	N	B49:01	0.0192
GEIVTEEHL	Both	1/4	4/4	9	P46821	44-52	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	B40:01	0.0322
RLMITDAARHK	IFN-α	0/4	1/4	11	P46821	125-135	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.0613
SMKPEPLFR	Both	1/4	2/4	9	P46821	392-400	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.2094
IVWHPANPAEK	IFN-α	0/4	1/4	11	P46821	473-483	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.3092
SLKPAAKPLPSK	Basal	3/4	0/4	12	P46821	544-555	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.1085
ATDVKPKAAK	Basal	1/4	0/4	10	P46821	625-634	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.2738
KSSTPLSEAK	Basal	1/4	0/4	10	P46821	741-750	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.0898
AAIGPAKEL	Both	3/4	4/4	9	P46821	815-823	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	C03:04	0.0066
SISAVSSEK	Basal	1/4	0/4	9	P46821	1242-1250	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.113
HLPTEVIEK	Both	1/4	2/4	9	P46821	1347-1355	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.6511
SESPIEKVL	IFN-α	0/4	4/4	9	P46821	1387-1395	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	B40:01	0.009
SLYQDKQEGK	Both	3/4	1/4	10	P46821	1452-1461	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.1829
SISEGTVSDK	Basal	1/4	0/4	10	P46821	1520-1529	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.2478
SISPPDFSPK	Both	3/4	1/4	10	P46821	1618-1627	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.0552
SISPPDFSPKTA	Basal	1/4	0/4	12	P46821	1618-1629	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	NA	NA
SLYASLTSEK	IFN-α	0/4	1/4	10	P46821	1760-1769	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.0189
FRASVLFDTM	IFN-α	0/4	1/4	10	P46821	1832-1841	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	C07:01	0.7701
LEKDSGGKTPGDF	Basal	1/4	0/4	13	P46821	1856-1868	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	NA	NA
SYAYQKPEETT	Basal	1/4	0/4	11	P46821	1869-1879	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	NA	NA
RTSDVGGYYY	IFN-α	0/4	1/4	10	P46821	1897-1906	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	C07:01/A03:01	1.0475/1.2398
SGYSYETIGK	Basal	1/4	0/4	10	P46821	1919-1928	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	1.1579
SEARQDVDL	IFN-α	0/4	2/4	9	P46821	2074-2082	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	B40:01	0.1651
ASPKPAGLK	Basal	1/4	0/4	9	P46821	2270-2278	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.5774
AEEPSRAV	IFN-α	0/4	2/4	8	P46821	2394-2401	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	B49:01	0.2957
AVLDALLEGK	Both	3/4	1/4	10	P46821;P78559	2400-2409	MAP1B;MAP1	MAP1L,MAP5,PP	Microtubule associated protein 1B;Microtubule	N	A03:01	0.089
IETEKTLYL	IFN-α	0/4	4/4	9	Q7KZI7;Q9P0L2;P	119-127	MARK2;MARK	EMK1,PAR-1,PAF	Microtubule affinity regulating kinase 2;Microtu	N	B40:01/B49:01	0.0894/0.1204
EEIDHAVHL	IFN-α	0/4	1/4	9	P53582	164-172	METAP1	KIAA0094,MAP1	Methionyl aminopeptidase 1	N	B40:01/B49:01	0.0522/0.1152
GEVDDGARKL	IFN-α	0/4	2/4	10	P53582	239-248	METAP1	KIAA0094,MAP1	Methionyl aminopeptidase 1	N	B40:01	0.0659
STFGNEEQPVLK	Basal	1/4	0/4	12	Q5JR59	883-894	MTUS2	CAZIP,ICIS,KIAA(	Microtubule associated tumor suppressor cand	N	A03:01	0.1256
ATSEIFASK	IFN-α	0/4	3/4	9	O75592	1076-1084	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	A03:01	0.0224
GEVLLRSI	IFN-α	0/4	3/4	8	O75592	1612-1619	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	B49:01	0.255
GEVLLRSIV	IFN-α	0/4	1/4	9	O75592	1612-1620	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	B49:01/B40:01	0.159/0.4635
KTHSGILGK	Both	1/4	4/4	9	O75592	2202-2210	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2; E3 ubiquitin protein ligas	N	A03:01	0.0068
YENYSFEEL	IFN-α	0/4	1/4	9	O75592	2379-2387	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	B40:01	0.0273
SEYDKQHSI	IFN-α	0/4	1/4	9	O75592	3482-3490	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	B49:01	0.0022
RLITSQVFGK	Basal	1/4	0/4	10	O75592	3903-3912	MYCBP2	FLJ10106,KIAA0	MYC binding protein 2, E3 ubiquitin protein ligas	N	A03:01	0.0475
GLFTHTIFY	IFN-α	0/4	3/4	9	Q9H3R1;O95803;I	477-485	NDST4;NDST3	HSST3,HSST,NST	N-deacetylase and N-sulfotransferase 4;N-deac	N	A03:01	0.0895
LIYEQVRLK	IFN-α	0/4	3/4	9	O76050	101-109	NEURL1	h-neu,neu-1,NEU	Neuralized E3 ubiquitin protein ligase 1	N	A03:01	0.0406
READDARLSV	IFN-α	0/4	1/4	10	O76050	241-250	NEURL1	h-neu,neu-1,NEU	Neuralized E3 ubiquitin protein ligase 1	N	B49:01	0.0788
RVAETIFVK	IFN-α	0/4	3/4	9	O76050	333-341	NEURL1	h-neu,neu-1,NEU	Neuralized E3 ubiquitin protein ligase 1	N	A03:01	0.0126
AETIFVKV	Both	4/4	4/4	8	O76050	335-342	NEURL1	NEURL,RNF67,n	Neuralized E3 ubiquitin protein ligase 1	N	B49:01	0.0924
GLHGTITQI	Both	1/4	2/4	9	O76050	435-443	NEURL1	NEURL,RNF67,n	Neuralized E3 ubiquitin protein ligase 1	N	A02:01	0.1674
ALSEILRSGK	Both	2/4	4/4	10	Q13562;Q15784	152-161	NEUROD1;NE	BETA2,BHF,MOD	Neuronal differentiation 2;Neuronal differentia	N	A03:01	0.1299
AEFEKNYAF	IFN-α	0/4	4/4	9	Q13562	291-299	NEUROD1	BETA2,BHF-1,bH	Neuronal differentiation 1	N	B40:01/B49:01	0.0397/0.0903
SVLEHHRDLMK	IFN-α	0/4	2/4	11	Q9BZQ4	97-107	NMNAT2	C1orf15,KIAA047	Nicotinamide nucleotide adenylyltransferase 2	N	A03:01	0.3103
FVFKPNSKK	Both	4/4	4/4	9	Q9UBD5	11-19	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	A03:01	0.0434
VFKPNSKK	Basal	2/4	0/4	8	Q9UBD5	12-19	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	NA	NA
KTDPKMLSK	Both	3/4	4/4	9	Q9UBD5	194-202	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	A03:01	0.0074
ATSPIIIHR	Both	3/4	3/4	9	Q9UBD5	253-261	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	A03:01	0.088
ATSPIIIHRL	IFN-α	0/4	1/4	10	Q9UBD5	253-262	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	E01:01	0.3748

Table S2, continued

KINEKVLQV	IFN-α	0/4	2/4	9	Q9UBD5	301-309	ORC3	IMAGE50150,LA	Origin recognition complex subunit 3	N	A02:01	0.0197
RTPSQIIRK	Both	4/4	3/4	9	Q01804	413-421	OTUD4	DUBA6,HSHIN1,I	OTU deubiquitinase 4	N	A03:01	0.0448
GEIIDIFKP	IFN-α	0/4	3/4	9	P19021	766-774	PAM	PAL,PHM	Peptidylglycine alpha-amidating monooxygenas	N	B49:01/B40:01	0.5047/1.3124
EEVESLKSI	IFN-α	0/4	2/4	9	Q9BZ23;Q9H999	240-248	PANK2;PANK3	C20orf48,FLJ117	Pantothenate kinase 2;Pantothenate kinase 3	N	B49:01	0.0553
SSFGNMMSK	IFN-α	0/4	1/4	9	Q9BZ23;Q8TE04	470-478	PANK2;PANK1	C20orf48,FLJ117	Pantothenate kinase 2;Pantothenate kinase 1	N	A03:01	0.0224
GEDSDEAQMRL	IFN-α	0/4	2/4	11	P26367	194-204	PAX6	AN,AN2,D11S81	Paired box 6	N	B40:01	0.194
GLISPGVSV	Basal	1/4	0/4	9	P26367	395-403	PAX6	AN,AN2,D11S81	Paired box 6	N	A02:01	0.0876
AEIPGGPEA	IFN-α	0/4	1/4	9	P29120	37-45	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin/kexin type 1	Y	B49:01	0.0656
GQIGSLENHYLFK	Both	1/4	1/4	13	P29120	59-71	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	NA	NA
YLFKHKNHPR	Both	4/4	4/4	10	P29120	68-77	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	NA	NA
SALRDSALNLFND	Both	4/4	2/4	13	P29120	111-123	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	NA	NA
RLAQKAFEY	IFN-α	0/4	2/4	9	P29120	282-290	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin/kexin type 1	Y	A03:01	0.1444
RRGDLHVTL	Both	4/4	4/4	9	P29120	517-525	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	C07:01	0.0165
IVNWKLILH	Both	1/4	3/4	9	P29120	584-592	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	A03:01	0.7197
RLLQSAFSK	Both	3/4	3/4	9	P29120	669-677	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	A03:01	0.0112
RLLQSAFSKNSPPK	Basal	2/4	0/4	14	P29120	669-682	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin/kexin type 1	Y	A03:01	1.3216
KLNKPSQLK	Both	2/4	2/4	9	P29120	706-714	PCSK1	NEC1,PC1,PC3,SI	Proprotein convertase subtilisin;kexin type 1	Y	A03:01	0.0032
LAETGAPRRF	Both	4/4	4/4	10	Q9UHG2	50-59	PCSK1N	BigLEN,PEN,SAA	Proprotein convertase subtilisin;kexin type 1 inl	Y	NA	NA
AEDQQARVL	IFN-α	0/4	3/4	9	Q9UHG2	98-106	PCSK1N	BigLEN,PEN,proS	Proprotein convertase subtilisin/kexin type 1 in	Y	B40:01	0.0197
PDAPAAQLARAL	Basal	1/4	0/4	12	Q9UHG2	129-140	PCSK1N	BigLEN,PEN,proS	Proprotein convertase subtilisin/kexin type 1 in	Y	NA	NA
SELPPEGVL	IFN-α	0/4	2/4	9	Q9UHG2	228-236	PCSK1N	BigLEN,PEN,proS	Proprotein convertase subtilisin/kexin type 1 in	Y	B40:01	0.0161
ERPVFTNHFLVE	IFN-α	0/4	1/4	12	P16519	26-37	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	NA	NA
FLVELHKGGEDKA	Basal	3/4	0/4	13	P16519	34-46	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	NA	NA
KLPFAEGLYH	Both	4/4	2/4	10	P16519	59-68	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin;kexin type 2	Y	A03:01	0.443
KLPFAEGLYHFYH	Basal	1/4	0/4	13	P16519	59-71	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	NA	NA
HFYHNGLA	IFN-α	0/4	1/4	8	P16519	68-75	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	NA	NA
HFYHNGLAK	Basal	1/4	0/4	9	P16519	68-76	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	A03:01	0.9278
HFYHNGLAKA	IFN-α	0/4	1/4	10	P16519	68-77	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	NA	NA
KTVDGPREL	Both	2/4	3/4	9	P16519	276-284	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin;kexin type 2	Y	C03:04	0.0747
RNGVGLEF	Both	4/4	3/4	8	P16519	431-438	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin;kexin type 2	Y	C07:01	1.4322
GEDARGTWTL	IFN-α	0/4	3/4	10	P16519	561-570	PCSK2	NEC2,PC2,SPC2	Proprotein convertase subtilisin/kexin type 2	Y	B40:01	0.0816
KEIDILPTQL	IFN-α	0/4	4/4	10	Q16654	70-79	PDK4		Pyruvate dehydrogenase kinase 4	N	B40:01	0.0257
KEDLTIKI	IFN-α	0/4	1/4	8	Q16654	284-291	PDK4		Pyruvate dehydrogenase kinase 4	N	B49:01	0.1198
KIRSDGTRYITK	IFN-α	0/4	4/4	12	Q9UPQ7;Q6ZMN7	924-935	PDZRN3;PDZR	KIAA1095,LNX3,S	PDZ domain containing ring finger 3;PDZ domai	N	A03:01	0.3694
KLTEENTTL	Both	3/4	4/4	9	Q86TG7	36-44	PEG10	HB,KIAA1051,ME	Paternally expressed 10	N	A02:01	0.0512
FEDPQRREV	Both	3/4	4/4	9	Q86TG7	158-166	PEG10	HB,KIAA1051,ME	Paternally expressed 10	N	B49:01	0.0589
LEVAKSLSA	IFN-α	0/4	1/4	9	Q86TG7	221-229	PEG10	HB-1,KIAA1051,M	Paternally expressed 10	N	B49:01	0.4126
HEYVAQNGI	IFN-α	0/4	4/4	9	Q86TG7	380-388	PEG10	HB-1,KIAA1051,M	Paternally expressed 10	N	B49:01	0.0186
RVYQPVRYY	Both	3/4	4/4	9	Q86TG7	501-509	PEG10	HB,KIAA1051,ME	Paternally expressed 10	N	A03:01	0.0138
IEMIPGAHSI	IFN-α	0/4	3/4	10	Q86TG7	533-542	PEG10	HB-1,KIAA1051,M	Paternally expressed 10	N	B49:01	0.0801
RVRQSPLATR	Both	4/4	4/4	10	O75381	40-49	PEX14		Peroxisomal biogenesis factor 14	N	A03:01	0.1376
ILFKDVIQA	IFN-α	0/4	1/4	9	Q15147	383-391	PLCB4		Phospholipase C beta 4	N	A02:01	0.0452
SLGRYRLLGK	IFN-α	0/4	1/4	10	Q9UL42	45-54	PNMA2	MA2,RGAG2	Paraneoplastic Ma antigen 2	N	A03:01	0.6907
RLLGKIFRK	Both	3/4	4/4	9	Q9UL42	50-58	PNMA2	MA2,RGAG2	Paraneoplastic Ma antigen 2	N	A03:01	0.006
RRTAQVRYL	Both	2/4	4/4	9	Q9UL42	247-255	PNMA2	MA2,RGAG2	Paraneoplastic Ma antigen 2	N	C07:01	0.0172
KEFETAETLL	IFN-α	0/4	2/4	10	O15514	24-33	POLR2D	RBP4	RNA polymerase II subunit D	N	B40:01	0.016
TEIQREIGI	IFN-α	0/4	4/4	9	O75145;Q13136;C	877-885	PPFIA3;PPFIA1	KIAA0654,LPNA3	PTPRF interacting protein alpha 3;PTPRF interac	N	B49:01	0.066
AEFSKTWEF	IFN-α	0/4	2/4	9	Q5SGD2	80-88	PPM1L	PP2CE	Protein phosphatase, Mg2+/Mn2+ dependent 1	N	B40:01/B49:01	0.09/0.1004
RLDEPHFGAK	Both	3/4	2/4	10	Q5SGD2	320-329	PPM1L	PP2CE	Protein phosphatase; Mg2;Mn2; dependent 1L	N	A03:01	0.0949
VEFNHSGELL	IFN-α	0/4	4/4	10	P63151	34-43	PPP2R2A	B55A,PR52A,PR5	Protein phosphatase 2 regulatory subunit Balph	N	B40:01/B49:01	0.1094/0.3115
LEIEEKINKI	IFN-α	0/4	4/4	10	P63151;Q00005;C	90-99	PPP2R2A;PPP2	B55A,PR52A,PR5	Protein phosphatase 2 regulatory subunit Balph	N	B49:01	0.2008
YETEHFNGV	IFN-α	0/4	1/4	9	Q16537;Q15172	230-238	PPP2R5E;PPP2	B56A,PR61A	Protein phosphatase 2 regulatory subunit B'eps	N	B49:01	0.0639

Table S2, continued

GEVDFKEFI	IFN-α	0/4	4/4	9	Q96LZ3;P63098	68-76	PPP3R2;PPP3I	PPP3RL,CALNB1,	Protein phosphatase 3 regulatory subunit B, bet	N	B49:01	0.0169
GELFQVLKM	IFN-α	0/4	1/4	9	Q96LZ3;P63098	110-118	PPP3R2;PPP3I	PPP3RL,CALNB1,	Protein phosphatase 3 regulatory subunit B, bet	N	B40:01/B49:01	0.2069/0.3065
TLLEKDYFL	Both	1/4	3/4	9	Q8TF05	892-900	PPP4R1	PP4R1	Protein phosphatase 4 regulatory subunit 1	N	A02:01	0.0229
KEIPLSEIL	IFN-α	0/4	2/4	9	Q15139;Q9BZL6	464-472	PRKD1;PRKD2	PKC-mu,PKCM,P	Protein kinase D1;Protein kinase D2	N	B40:01	0.0051
RMWEIAIQH	IFN-α	0/4	4/4	9	Q15139	530-538	PRKD1	PKC-mu,PKCM,P	Protein kinase D1	N	A03:01	0.1274
RLPQPPVGK	Both	1/4	2/4	9	Q16849	156-164	PTPRN	IA	Protein tyrosine phosphatase; receptor type N	Y	A03:01	0.0088
GEKPASPAV	Both	2/4	3/4	9	Q16849	321-329	PTPRN	IA	Protein tyrosine phosphatase; receptor type N	Y	B49:01	0.0277
LEILAEHV	IFN-α	0/4	2/4	8	Q16849	490-497	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	B49:01	0.3143
VVGPALTF	IFN-α	0/4	2/4	8	Q16849;202	509-516	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	NA	NA
VVGPALTFR	IFN-α	0/4	1/4	9	Q16849;202	509-517	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	A03:01	0.7072
SLADVTQQA	IFN-α	0/4	1/4	9	Q16849	526-534	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	A02:01	0.0469
SLYHVYEV	Both	2/4	4/4	8	Q16849	830-837	PTPRN	IA	Protein tyrosine phosphatase; receptor type N	Y	A02:01	0.2596
KEIDIAATL	IFN-α	0/4	4/4	9	Q92932;Q16849	970-978	PTPRN2;PTPR	IA-2,IA-2beta,ICA	Protein tyrosine phosphatase, receptor type N;	Y	B40:01	0.0032
KEIDIAATLE	IFN-α	0/4	2/4	10	Q92932;Q16849	970-979	PTPRN2;PTPR	IA-2,IA-2beta,ICA	Protein tyrosine phosphatase, receptor type N;	Y	B40:01	0.3669
AVAEEVNAILK	Both	4/4	4/4	11	Q92932;Q16849	1001-1011	PTPRN2;PTPR	IA,IA,2beta,ICAA	Protein tyrosine phosphatase; receptor type N2	Y	A03:01	0.4261
RLYQEVHRL	IFN-α	0/4	3/4	9	Q92932	387-395	PTPRN2	IA-2beta,ICAAR,I	Protein tyrosine phosphatase, receptor type N2	N	A02:01	0.0126
VLGPAVTFK	IFN-α	0/4	1/4	9	Q92932	549-557	PTPRN2	IA-2beta,ICAAR,I	Protein tyrosine phosphatase, receptor type N2	N	A03:01	0.0148
RLIVIGDSTVGK	Basal	1/4	0/4	12	Q96DA2;Q14964	10-21	RAB39B;RAB3	MRX72,RAB39	RAB39B, member RAS oncogene family;RAB39A	N	A03:01	0.092
IVIGDSTVGK	Basal	1/4	0/4	10	Q96DA2;Q14964	12-21	RAB39B;RAB3	MRX72,RAB39	RAB39B, member RAS oncogene family;RAB39A	N	A03:01	0.2225
KLLIIGNSSVGK	Both	1/4	1/4	12	P20337;Q96E17	24-35	RAB3B;RAB3C		RAB3C; member RAS oncogene family;RAB3B; r	N	A03:01	0.2813
LLIIGNSSVGK	Basal	1/4	0/4	11	P20337;Q96E17	25-35	RAB3B;RAB3C		RAB3B, member RAS oncogene family;RAB3C, r	N	A03:01	0.9429
LIIGNSSVGK	Basal	1/4	0/4	10	P20337;Q96E17;P	26-35	RAB3B;RAB3C;RAB3A		RAB3B, member RAS oncogene family;RAB3C, r	N	A03:01	0.8175
IIGNSSVGK	Both	4/4	2/4	9	P20336;P20337;Q	27-35	RAB3A;RAB3B;RAB3C		RAB3C; member RAS oncogene family;RAB3B; r	N	A03:01	0.3243
IIGNSSVGKTSF	Basal	1/4	0/4	12	P20337;Q96E17;P	27-38	RAB3B;RAB3C;RAB3A		RAB3B, member RAS oncogene family;RAB3C, r	N	NA	NA
LFRYADDTFTPA	Basal	1/4	0/4	12	P20337	39-50	RAB3B		RAB3B, member RAS oncogene family	N	NA	NA
RYADDTFTPA	Basal	1/4	0/4	10	P20337	41-50	RAB3B		RAB3B, member RAS oncogene family	N	NA	NA
IWDTAGQERY	Basal	1/4	0/4	10	P20337;Q96E17;P	75-84	RAB3B;RAB3C	D2-2,GOV,RAB1(	RAB3B, member RAS oncogene family;RAB3C, r	N	NA	NA
KTYSWDNA	IFN-α	0/4	1/4	8	P20337;O95716;C	121-128	RAB3D;RAB3C	D2-2,GOV,RAB1(	RAB3D, member RAS oncogene family;RAB3C, r	N	NA	NA
RLSDTPPLL	IFN-α	0/4	3/4	9	P20337	205-213	RAB3B		RAB3B, member RAS oncogene family	N	A02:01	0.023
KLFPGSPAI	IFN-α	0/4	2/4	9	P49792;Q99666;C	123-131	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A02:01	0.006
KLFPGSPAIYK	IFN-α	0/4	3/4	11	P49792;Q99666;C	123-133	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.0023
HEAERNIAL	IFN-α	0/4	4/4	9	P49792;Q99666;C	189-197	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	B40:01	0.0162
VVYTSHLQLK	IFN-α	0/4	4/4	10	P49792;Q99666;C	491-500	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.027
KVLPLLKIIKK	IFN-α	0/4	2/4	11	P49792;Q99666;C	608-618	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.0492
YEEDAHITF	IFN-α	0/4	3/4	9	P49792;Q99666;C	646-654	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	B40:01/B49:01	0.0675/0.1409
SLSPSKSYK	Basal	1/4	0/4	9	P49792;Q7Z3J3;O	786-794	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.0109
LLRPAANVTPTK	Basal	2/4	0/4	12	P49792;Q7Z3J3;O	888-899	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.8285
GLKDFKTFL	IFN-α	0/4	2/4	9	P49792;Q99666;C	2176-2184	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A02:01	0.381
KLQDVADSFKK	IFN-α	0/4	1/4	11	P49792;Q99666;C	2424-2434	RANBP2;RGPI	ADANE,ANE1,NU	RAN binding protein 2;RANBP2-like and GRIP do	N	A03:01	0.1945
AEGIEYLRV	IFN-α	0/4	1/4	9	Q684P5	188-196	RAP1GAP2	GARNL4,KIAA10	RAP1 GTPase activating protein 2	N	B49:01	0.1691
KLPSVPQIAK	Both	3/4	3/4	10	Q684P5	216-225	RAP1GAP2	GARNL4,KIAA10	RAP1 GTPase activating protein 2	N	A03:01	0.0291
ALLDNLHDEL	IFN-α	0/4	1/4	10	Q684P5	459-468	RAP1GAP2	GARNL4,KIAA10	RAP1 GTPase activating protein 2	N	A02:01	0.0559
VLHKKALRNK	Both	3/4	2/4	10	Q9Y272	216-225	RASD1	AGS1,DEXRAS1	Ras related dexamethasone induced 1	N	A03:01	0.3379
KLDEYIAIA	IFN-α	0/4	1/4	9	Q9UKL0;Q9P2K3	148-156	RCOR1;RCOR	COREST,KIAA007	REST corepressor 1;REST corepressor 3	N	A02:01	0.1025
KLDEYIAIAK	IFN-α	0/4	4/4	10	Q9UKL0;Q9P2K3	148-157	RCOR1;RCOR	COREST,KIAA007	REST corepressor 1;REST corepressor 3	N	A03:01	0.2511
KTITQIVKDK	Both	3/4	4/4	10	Q8HWS3	108-117	RFX6	MGC33442,RFXD	Regulatory factor X6	N	A03:01	0.0871
ATFGKTIRQK	Both	4/4	4/4	10	Q8HWS3	164-173	RFX6	MGC33442,RFXD	Regulatory factor X6	N	A03:01	0.0066
FGKTIRQK	IFN-α	0/4	1/4	8	Q8HWS3	166-173	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	NA	NA
IGIKESSAYYHSVY	IFN-α	0/4	1/4	14	Q8HWS3	195-208	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	NA	NA
SVYSGKGLTR	Both	4/4	4/4	10	Q8HWS3	206-215	RFX6	MGC33442,RFXD	Regulatory factor X6	N	A03:01	0.086
SVYSGKGLTRF	Both	3/4	3/4	11	Q8HWS3	206-216	RFX6	MGC33442,RFXD	Regulatory factor X6	N	A03:01	1.7177
YSGKGLTR	IFN-α	0/4	2/4	8	Q8HWS3	208-215	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	NA	NA

Table S2, continued

KLKNEGGFTR	Basal	1/4	0/4	10	Q8HWS3	220-229	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	A03:01	0.6272
KLKNEGGFTRK	IFN-α	0/4	1/4	11	Q8HWS3	220-230	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	A03:01	0.1088
FEEIQHFLL	IFN-α	0/4	2/4	9	Q8HWS3	284-292	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	B40:01	0.1311
SMVSDIERV	Both	4/4	4/4	9	Q8HWS3	408-416	RFX6	MGC33442,RFXI	Regulatory factor X6	N	A02:01	0.0681
VVEQRVIKTSK	Both	2/4	4/4	11	Q8HWS3	472-482	RFX6	MGC33442,RFXI	Regulatory factor X6	N	A03:01	1.8733
KEQELQNLL	IFN-α	0/4	4/4	9	Q8HWS3	544-552	RFX6	dJ955L16.1,MGC	Regulatory factor X6	N	B40:01	0.0107
GLYPHHTEH	Both	3/4	4/4	9	Q8HWS3	716-724	RFX6	MGC33442,RFXI	Regulatory factor X6	N	A03:01	0.1466
KLFPGSPAIY	Both	4/4	4/4	10	Q99666;P0DJD0;P	123-132	RGPD5;RGPD3	ADANE,ANE1,BS	RANBP2;like and GRIP domain containing 5;RAN	N	A03:01	0.0969
VLMEQIKLL	IFN-α	0/4	1/4	9	Q99666;O14715;A	1681-1689	RGPD5;RGPD8	BS-63,DKFZp686	RANBP2-like and GRIP domain containing 5;RAN	N	A02:01	0.0215
GRVRTKTVKKAA	Basal	3/4	0/4	12	P08708	2-13	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
KAARVIIEK	Both	3/4	2/4	9	P08708	11-19	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	A03:01	0.0648
YTRLGNDF	Basal	1/4	0/4	8	P08708	21-28	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
YTRLGNDFHTNK	Basal	2/4	0/4	12	P08708	21-32	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
RLGNDFHTNK	Both	4/4	4/4	10	P08708	23-32	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	A03:01	0.0299
EEIAIIPSK	Basal	1/4	0/4	9	P08708	36-44	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
IAIIPSKKL	Both	2/4	4/4	9	P08708	38-46	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	C03:04	0.0272
AIIPSKKLRNK	Both	4/4	4/4	11	P08708	39-49	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	A03:01	0.1749
IEVDPDTKEML	IFN-α	0/4	3/4	11	P08708	96-106	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	B40:01	0.1571
PDTKEMLKLL	Basal	1/4	0/4	10	P08708	100-109	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
LKLLDFGSLS	IFN-α	0/4	1/4	10	P08708	106-115	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
KLLDFGSLSNL	Both	3/4	3/4	11	P08708	107-117	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	A02:01	0.2199
NLQVTQPTVGMNF	IFN-α	0/4	2/4	13	P08708	116-128	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
LQVTQPTVGMNF	IFN-α	0/4	2/4	12	P08708	117-128	RPS17	MGC72007,RPS1	Ribosomal protein S17	N	NA	NA
STYFTGILQK	Both	4/4	4/4	10	Q16799	120-129	RTN1	NSP	Reticulon 1	N	A03:01	0.0091
RIYKSVLQAVQK	Both	2/4	3/4	12	Q16799	639-650	RTN1	NSP	Reticulon 1	N	A03:01	0.0406
AILEQILSH	Both	4/4	1/4	9	Q96NL0;Q59EK9	68-76	RUNDC3A;RU	RAP2IP,RPIB9,RF	RUN domain containing 3A;RUN domain contai	N	A03:01	0.2811
KLIGDTPIDTFL	IFN-α	0/4	3/4	12	P19793;P48443;P	440-451	RXRA;RXRG;R	NR2B1,NR2B3,H	Retinoid X receptor alpha;Retinoid X receptor g	N	A02:01	0.4879
GVSVPLQQK	Both	1/4	2/4	9	P13521	83-91	SCG2	CHGC,SN,SgII	Secretogranin II	Y	A03:01	0.0443
YEENSRDNPF	Both	3/4	2/4	10	P13521	170-179	SCG2	CHGC,SN,SgII	Secretogranin II	Y	B40:01	1.2646
ATLESVFQELGK	Basal	1/4	0/4	12	P13521	196-207	SCG2	CHGC,SgII,SN	Secretogranin II	Y	A03:01	0.5238
LVNAAGSGR	Basal	2/4	0/4	9	P13521	317-325	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA
KTSYFPNPY	IFN-α	0/4	1/4	9	P13521	449-457	SCG2	CHGC,SgII,SN	Secretogranin II	Y	A03:01/C07:01	0.9091/1.4088
TSYFPNPYNQEK	Basal	1/4	0/4	12	P13521	450-461	SCG2	CHGC,SgII,SN	Secretogranin II	Y	A03:01	1.3754
NQEKVLPRLP	Basal	1/4	0/4	10	P13521	458-467	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA
NQEKVLPRLPY	Basal	1/4	0/4	11	P13521	458-468	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA
VKYPEIINSNQV	Basal	1/4	0/4	12	P13521	513-524	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA
ELSAERPLNEQIAE	Basal	1/4	0/4	14	Q8WXD2	35-48	SCG3	FLJ90833,SgIII	Secretogranin III	Y	NA	NA
AERPLNEQI	IFN-α	0/4	2/4	9	Q8WXD2	38-46	SCG3	FLJ90833,SgIII	Secretogranin III	Y	B49:01	0.0194
SFVDNLNLLK	Basal	2/4	0/4	10	Q8WXD2	70-79	SCG3	FLJ90833,SgIII	Secretogranin III	Y	A03:01	0.8885
ITEKEKIEKE	Basal	1/4	0/4	10	Q8WXD2	81-90	SCG3	FLJ90833,SgIII	Secretogranin III	Y	NA	NA
KLIDDYDSTK	Both	3/4	4/4	10	Q8WXD2	115-124	SCG3	FLJ90833,SgIII	Secretogranin III	Y	A03:01	0.2186
AEDIVHKI	IFN-α	0/4	4/4	8	Q8WXD2	148-155	SCG3	FLJ90833,SgIII	Secretogranin III	Y	B49:01	0.032
AVFDKIVSK	Both	3/4	4/4	9	Q8WXD2	166-174	SCG3	FLJ90833,SgIII	Secretogranin III	Y	A03:01	0.0022
TKTYSEDNFEELQ	Basal	1/4	0/4	13	Q8WXD2	263-275	SCG3	FLJ90833,SgIII	Secretogranin III	Y	NA	NA
KEKETLITI	IFN-α	0/4	2/4	9	Q8WXD2	294-302	SCG3	FLJ90833,SgIII	Secretogranin III	Y	B49:01	0.0067
KLFPAPSEK	Both	2/4	1/4	9	Q8WXD2	353-361	SCG3	FLJ90833,SgIII	Secretogranin III	Y	A03:01	0.0014
RVSEADIQRLLH	Basal	1/4	0/4	12	P05408	34-45	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SEADIQRLL	IFN-α	0/4	3/4	9	P05408	36-44	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B40:01/B49:01	0.0104/0.0188
AEFSREFQL	IFN-α	0/4	4/4	9	P05408	138-146	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B40:01/B49:01	0.0156/0.0449
REFQLHQHL	IFN-α	0/4	4/4	9	P05408	142-150	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B40:01	0.008
HLFDPEHDYPG	Both	3/4	4/4	11	P05408	149-159	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
HLFDPEHDYPGL	Basal	1/4	0/4	12	P05408	149-160	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A02:01	0.4226

Table S2, continued

HLFDPEHDYPGLG	Both	4/4	4/4	13	P05408	149-161	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
HLFDPEHDYPGLGK	Basal	3/4	0/4	14	P05408	149-162	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	0.5414
LFDPEHDYPG	Basal	1/4	0/4	10	P05408	150-159	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
LFDPEHDYPGL	Both	3/4	4/4	11	P05408	150-160	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
LFDPEHDYPGLG	Basal	1/4	0/4	12	P05408	150-161	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
LFDPEHDYPGLGKW	Both	1/4	1/4	14	P05408	150-163	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
FDPEHDYPG	Basal	1/4	0/4	9	P05408	151-159	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
FDPEHDYPGL	Basal	1/4	0/4	10	P05408	151-160	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
FDPEHDYPGLGKW	Basal	1/4	0/4	13	P05408	151-163	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
HDYPGLGKW	Basal	1/4	0/4	9	P05408	155-163	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B49:01	1.3213
NKKLLYEKMKGGE	Basal	1/4	0/4	13	P05408	164-176	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVNPYLQGQR	Both	4/4	4/4	10	P05408	182-191	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.4187
SVNPYLQGQRL	Both	4/4	4/4	11	P05408	182-192	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVNPYLQGQRLD	Both	4/4	4/4	12	P05408	182-193	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVNPYLQGQRLDN	Both	4/4	4/4	13	P05408	182-194	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVNPYLQGQRLDNV	Basal	4/4	0/4	14	P05408	182-195	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
YLQGQRLDNV	Basal	3/4	0/4	10	P05408	186-195	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A02:01	0.6607
YLQGQRLDNVV	Basal	2/4	0/4	11	P05408	186-196	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A02:01	1.9699
YLQGQRLDNVVA	Both	4/4	4/4	12	P05408	186-197	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
RLDNVVAKK	Both	3/4	4/4	9	P05408	191-199	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	0.0403
VVAKKSVPH	Both	4/4	4/4	9	P05408	195-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.0522
VAKKSVPH	IFN-α	0/4	3/4	8	P05408	196-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVPHFSDEDKD	Basal	1/4	0/4	11	P05408	200-210	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
SVPHFSDEDKDPE	Basal	1/4	0/4	13	P05408	200-212	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
HFSDEDKDPE	Basal	1/4	0/4	10	P05408	203-212	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
FSDEDKDPE	Basal	1/4	0/4	9	P05408	204-212	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA
KELDAFFL	IFN-α	0/4	3/4	8	O76038	35-42	SCGN	CALBL,DJ501N12	Secretagogin, EF-hand calcium binding protein	N	B40:01/B49:01	0.4983/0.7761
RIKPPGPTAVPK	Both	3/4	4/4	12	Q9UQR0	285-296	SCML2		Sex comb on midleg;like 2 (Drosophila)	N	A03:01	0.0415
SVKNITPRKK	IFN-α	0/4	1/4	10	Q9UQR0	300-309	SCML2		Sex comb on midleg-like 2 (Drosophila)	N	A03:01	0.2071
IAVPDPSVL	IFN-α	0/4	1/4	9	Q9UQR0	612-620	SCML2		Sex comb on midleg-like 2 (Drosophila)	N	C03:04	0.0023
HEIDGKALFL	IFN-α	0/4	4/4	10	Q9UQR0	658-667	SCML2		Sex comb on midleg-like 2 (Drosophila)	N	B40:01	0.0484
KLGPALKL	IFN-α	0/4	2/4	8	Q9UQR0;Q96GD3	680-687	SCML2;SCMH	Scml3	Sex comb on midleg-like 2 (Drosophila);Sex con	N	E01:01	1.5466
GEGLFSKV	IFN-α	0/4	4/4	8	Q8WTS6	227-234	SETD7	KIAA1717,KMT7,	SET domain containing lysine methyltransferase	N	B49:01	0.4426
KAFQATQQK	Both	4/4	4/4	9	Q8WTS6	358-366	SETD7	KIAA1717,KMT7,	SET domain containing lysine methyltransferase	N	A03:01	0.0732
RLFKEREKPLK	Both	1/4	2/4	11	O43166	273-283	SIPA1L1	E6TP1,KIAA0440	Signal induced proliferation associated 1 like 1	N	A03:01	0.1402
SSTYPSLPK	Basal	2/4	0/4	9	O43166	1616-1624	SIPA1L1	E6TP1,KIAA0440	Signal induced proliferation associated 1 like 1	N	A03:01	0.2655
KIAPESAIK	Both	1/4	3/4	9	Q6KCM7;Q9BV35	252-260	SLC25A25;SLC	APC2,FLJ30339,K	Solute carrier family 25 member 25;Solute carri	N	A03:01	0.0389
RTYLVNDKAAK	Both	4/4	4/4	11	Q8IWU4	6-16	SLC30A8	ZNT8,ZnT	Solute carrier family 30 member 8	Y	A03:01	0.1627
RTYLVNDKAAKMY	Basal	4/4	0/4	13	Q8IWU4	6-18	SLC30A8	ZnT-8,ZNT8	Solute carrier family 30 member 8	Y	NA	NA
KMYAFTLE	IFN-α	0/4	3/4	8	Q8IWU4	16-23	SLC30A8	ZnT-8,ZNT8	Solute carrier family 30 member 8	Y	NA	NA
RVQDIFTAGK	IFN-α	0/4	3/4	10	Q9UHI5	185-194	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	A03:01	0.0385
AVAVTFGEK	IFN-α	0/4	1/4	9	Q9UHI5;Q9NS82	301-309	SLC7A8;SLC7A	LAT2,LPI-PC1,asc	Solute carrier family 7 member 8;Solute carrier	N	A03:01	0.2709
VTFGEKLLGV	IFN-α	0/4	2/4	10	Q9UHI5	304-313	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	A02:01	1.4367
REGHLPSVL	IFN-α	0/4	4/4	9	Q9UHI5	346-354	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	B40:01	0.0594
GVYWQHKPK	IFN-α	0/4	2/4	9	Q9UHI5	466-474	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	A03:01	0.0912
SLFSFTSPAVK	Both	3/4	3/4	11	Q15797;O15198;C	5-15	SMAD1;SMAD	Dwfc,JV4,JV5,MA	SMAD family member 1;SMAD family member 5	N	A03:01	0.085
FSFTSPAV	IFN-α	0/4	1/4	8	Q15797;Q99717;C	7-14	SMAD1;SMAD	JV4-1,MADH1,M	SMAD family member 1;SMAD family member 5	N	C03:04	0.8525
VLYVGSKTK	Both	2/4	4/4	9	P37840	37-45	SNCA	NACP,PARK1,PAI	Synuclein alpha	N	A03:01	0.0309
GVVHGVATV	IFN-α	0/4	2/4	9	P37840	47-55	SNCA	alpha-synuclein,/	Synuclein alpha	N	A02:01	0.4519
SEELMTIKL	IFN-α	0/4	1/4	9	O60284	906-914	ST18	KIAA0535,NZF3,/	ST18, C2H2C-type zinc finger	N	B40:01	0.0423
SSYVGRGL	IFN-α	0/4	4/4	8	P52630;Q14765	505-512	STAT2;STAT4	STAT113,	Signal transducer and activator of transcription	N	C03:04	0.8759
RASGQAFEL	Both	2/4	4/4	9	Q93045;P16949	48-56	STMN2;STMN	C1orf215,FLJ322	Stathmin 1;Stathmin 2	N	C03:04	0.1447

Table S2, continued

KLEAAEERRK	Both	4/4	4/4	10	Q93045;P16949	87-96	STMN2;STMN	C1orf215,FLJ322	Stathmin 4;Stathmin 1;Stathmin 2	N	A03:01	1.188
RLEAAEERRK	IFN-α	0/4	3/4	10	Q9NZ72	88-97	STMN3	SCLIP	Stathmin 3	N	A03:01	1.0569
KTQEAQVLK	Both	2/4	4/4	9	Q9NZ72	97-105	STMN3	SCLIP	Stathmin 3	N	A03:01	0.0384
AEEKLNYKM	IFN-α	0/4	2/4	9	Q9NZ72	132-140	STMN3	SCLIP	Stathmin 3	N	B40:01	0.1098
REKELHAAEV	IFN-α	0/4	1/4	10	Q9NZ72	159-168	STMN3	SCLIP	Stathmin 3	N	B49:01/B40:01	0.3997/0.9464
GEKIMHDVI	IFN-α	0/4	1/4	9	P61764	11-19	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B49:01/B40:01	0.3827/0.5542
KIMHDVIKK	Both	3/4	3/4	9	P61764	13-21	STXBP1	MUNC18,UNC18	Syntaxin binding protein 1	N	A03:01	0.0154
GEWKVLVV	IFN-α	0/4	4/4	8	P61764	26-33	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B49:01	0.1853
SEKSVHSLI	IFN-α	0/4	2/4	9	P61764	80-88	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B49:01	0.0365
RLAEQIATL	Basal	1/4	0/4	9	P61764	171-179	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	A02:01	0.0108
GEYKDNAL	IFN-α	0/4	3/4	8	P61764	193-200	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B40:01	0.1783
GEYKDNALL	Both	3/4	2/4	9	P61764	193-201	STXBP1	MUNC18,UNC18	Syntaxin binding protein 1	N	B40:01	0.016
GEYKDNALLA	IFN-α	0/4	4/4	10	P61764	193-202	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B49:01/B40:01	0.3787/1.0777
HIAEVSQEV	IFN-α	0/4	2/4	9	P61764	295-303	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	A02:01	0.0813
VSQEVTRSL	IFN-α	0/4	3/4	9	P61764	299-307	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	C03:04	0.0348
SLKDFSSSK	Both	3/4	4/4	9	P61764	306-314	STXBP1	MUNC18,UNC18	Syntaxin binding protein 1	N	A03:01	0.0208
TTMRDLSQMLK	Both	1/4	2/4	11	P61764	322-332	STXBP1	MUNC18,UNC18	Syntaxin binding protein 1	N	A03:01	0.5332
AHLGVPIVTDSTL	Both	2/4	1/4	13	P61764	444-456	STXBP1	MUNC18,UNC18	Syntaxin binding protein 1	N	NA	NA
RISEQTYQL	IFN-α	0/4	3/4	9	P61764	467-475	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	A02:01	0.1205
GEYRSGPRLI	IFN-α	0/4	2/4	10	P61764	529-538	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	B49:01	0.0993
IGSTHILTPQK	Basal	2/4	0/4	11	P61764	567-577	STXBP1	hUNC18,MUNC1	Syntaxin binding protein 1	N	NA	NA
GEYQGIPRA	IFN-α	0/4	4/4	9	Q7L0J3	96-104	SV2A	KIAA0736,SV2	Synaptic vesicle glycoprotein 2A	N	B49:01	0.0527
AVDYASRTK	Both	2/4	1/4	9	Q7L0J3	477-485	SV2A	KIAA0736,SV2	Synaptic vesicle glycoprotein 2A	N	A03:01	0.2089
KISDFGLSK	IFN-α	0/4	1/4	9	P43405;P43403;Q	509-517	SYK;ZAP70;AN	SRK,STD,ZAP-70,	Spleen associated tyrosine kinase;Zeta chain of	N	A03:01	0.0077
KEPLGFVKV	IFN-α	0/4	2/4	9	P08247	23-31	SYP	MRX96	Synaptophysin	N	B49:01	0.1302
AIKLPAGGK	Basal	1/4	0/4	9	O43581	70-78	SYT7	IPCA-7,MGC150	Synaptotagmin 7	N	A03:01	0.241
KQMEQISQFL	IFN-α	0/4	4/4	10	P37802;Q9UI15	88-97	TAGLN2;TAGL	HA1756,KIAA012	Transgelin 2;Transgelin 3	N	A02:01	0.3529
TDIFQTVDL	IFN-α	0/4	1/4	9	P37802;Q9UI15	108-116	TAGLN2;TAGL	HA1756,KIAA012	Transgelin 2;Transgelin 3	N	NA	NA
AAAEAGLHV	IFN-α	0/4	2/4	9	Q13207	69-77	TBX2		T-box 2	N	C03:04	0.3759
ALGPHPPAA	IFN-α	0/4	1/4	9	Q13207	79-87	TBX2		T-box 2	N	A02:01	0.1096
VEDDPKVTL	IFN-α	0/4	4/4	9	Q13207	101-109	TBX2		T-box 2	N	B40:01	0.0086
ELWDQFHKL	IFN-α	0/4	2/4	9	Q13207	113-121	TBX2		T-box 2	N	A02:01	0.0947
YQNDKITQL	IFN-α	0/4	1/4	9	Q13207;O15119	266-274	TBX2;TBX3	TBX3-ISO,UMS,X	T-box 2;T-box 3	N	C07:01	0.0175
SEERAGAPL	Both	1/4	1/4	9	Q13207	368-376	TBX2		box 2	N	B40:01	0.177
RLNKAILRY	Both	1/4	3/4	9	P35443	94-102	THBS4		Thrombospondin 4	N	A03:01	0.0358
RLAATQVEY	Both	3/4	4/4	9	Q9BTF0	37-45	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	A03:01	0.1781
KQFPLIISSV	IFN-α	0/4	1/4	10	Q9BTF0	76-85	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	A02:01	0.3711
LELDAKKEKL	IFN-α	0/4	2/4	10	Q9BTF0	116-125	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	B40:01	0.2478
QEVGKVIGI	IFN-α	0/4	4/4	9	Q9BTF0	224-232	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	B49:01	0.0419
KVIGIAIMK	Both	2/4	3/4	9	Q9BTF0	228-236	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	A03:01	0.0117
KEWPDVYYV	IFN-α	0/4	4/4	9	Q9BTF0	318-326	THUMPD2	C2orf8,MGC245	THUMP domain containing 2	N	B49:01	0.0292
RSLSEDVRFYY	Both	2/4	4/4	11	Q92545	709-719	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	NA	NA
SEDVRFYY	IFN-α	0/4	2/4	8	Q92545	712-719	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	NA	NA
AIFEVNTDLQK	IFN-α	0/4	1/4	11	Q92545	804-814	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	A03:01	0.1143
KLREPNFTLK	IFN-α	0/4	4/4	10	Q92545	1007-1016	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	A03:01	0.008
RLSPAPLAH	IFN-α	0/4	1/4	9	Q92545	1320-1328	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	A03:01	0.0464
GEVFSKLGL	IFN-α	0/4	1/4	9	Q92545	1735-1743	TMEM131	CC28,KIAA0257,I	Transmembrane protein 131	N	B40:01	0.0775
SLDEKVVSI	IFN-α	0/4	2/4	9	Q14C87	737-745	TMEM132D	KIAA1944,MOLT	Transmembrane protein 132D	N	A02:01	0.0035
AIKKDFTNK	IFN-α	0/4	1/4	9	Q86VY9	449-457	TMEM200A	KIAA1913,TTMC	Transmembrane protein 200A	N	A03:01	0.0781
ALLPAGLRQK	IFN-α	0/4	2/4	10	P28289	42-51	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A03:01	0.0126
KEFKDREDL	IFN-α	0/4	4/4	9	P28289	76-84	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	B40:01	0.0613
AAILGMHTL	IFN-α	0/4	1/4	9	P28289	132-140	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	C03:04	0.0202

Table S2, continued

ALKENSYVK	Both	3/4	4/4	9	P28289	219-227	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A03:01	0.1766
ALKENSYVKK	Both	3/4	4/4	10	P28289	219-228	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A03:01	0.127
AEMLKENKV	IFN-α	0/4	2/4	9	P28289	245-253	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	B49:01	0.0418
FISGAGILRL	IFN-α	0/4	1/4	10	P28289	263-272	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A02:01/C03:04	1.1785/1.9369
RLADLTGPIIPK	Both	2/4	4/4	12	P28289	343-354	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A03:01	0.0432
LTGPIIPK	Both	4/4	3/4	8	P28289	347-354	TMOD1	D9S57E,ETMOD,	Tropomodulin 1	N	A03:01	1.5623
PEDPHPLSVV	IFN-α	0/4	1/4	10	O75382	99-108	TRIM3	BERP,HAC1,RNF	Tripartite motif containing 3	N	B49:01	1.8937
STSPYFLEH	Both	3/4	4/4	9	P41732	160-168	TSPAN7	A15,CD231,DXS1	Tetraspanin 7	Y	A03:01/C07:01	0.3166/0.7918
MASHRLLLL	IFN-α	0/4	1/4	9	P02766	9-21	TTR	CTS,CTS1,HsT26	Transthyretin	N	C03:04	0.1675
YSYSTTAVVTNPK	Basal	1/4	0/4	13	P02766	134-146	TTR	CTS,CTS1,HsT26	Transthyretin	N	NA	NA
SYSTTAVVTNPKE	Basal	1/4	0/4	13	P02766	135-147	TTR	CTS,CTS1,HsT26	Transthyretin	N	NA	NA
STTAVVTNPK	Both	4/4	1/4	10	P02766	137-146	TTR	CTS,CTS1,HsT26	Transthyretin	N	A03:01	0.577
STTAVVTNPKE	Basal	1/4	0/4	11	P02766	137-147	TTR	CTS,CTS1,HsT26	Transthyretin	N	NA	NA
TTAVVTNPK	Both	4/4	3/4	9	P02766	138-146	TTR	CTS,CTS1,HsT26	Transthyretin	N	A03:01	0.5119
IEDGKGTILI	IFN-α	0/4	4/4	10	Q9UBT2	482-491	UBA2	ARX,FLJ13058,Hl	Ubiquitin like modifier activating enzyme 2	N	B49:01/B40:01	0.4859/0.7083
KLDEKENLSAK	Both	3/4	1/4	11	Q9UBT2	613-623	UBA2	ARX,FLJ13058,Hl	Ubiquitin like modifier activating enzyme 2	N	A03:01	0.1735
IEQKEELDDVIA	Basal	1/4	0/4	12	Q9UBT2	627-638	UBA2	ARX,FLJ13058,Hl	Ubiquitin like modifier activating enzyme 2	N	NA	NA
IEQKEELDDVIALD	Basal	1/4	0/4	14	Q9UBT2	627-640	UBA2	ARX,FLJ13058,Hl	Ubiquitin like modifier activating enzyme 2	N	NA	NA
GTIGLIHAV	Both	4/4	4/4	9	P09936	91-99	UCHL1	PARK5,PGP9,Uch	Ubiquitin C;terminal hydrolase L1	N	A02:01	0.1773
KQFLSETEK	Basal	1/4	0/4	9	P09936	115-123	UCHL1	PARK5,PGP9.5,U	Ubiquitin C-terminal hydrolase L1	N	A03:01	0.3288
GEVRFSAV	Both	2/4	4/4	8	P09936	210-217	UCHL1	PARK5,PGP9,Uch	Ubiquitin C;terminal hydrolase L1	N	B49:01	0.541
SEAEKGQWEDASL	IFN-α	0/4	3/4	13	Q969E3	41-53	UCN3	SPC,UCNIII	Urocortin 3	Y	B40:01	0.5808
SEAEKGQWEDASLL	IFN-α	0/4	2/4	14	Q969E3	41-54	UCN3	SPC,UCNIII	Urocortin 3	Y	B40:01	1.1068
AEKGQWEDASL	IFN-α	0/4	3/4	11	Q969E3	43-53	UCN3	SPC,UCNIII	Urocortin 3	Y	B40:01	0.3317
AEKGQWEDASLL	IFN-α	0/4	2/4	12	Q969E3	43-54	UCN3	SPC,UCNIII	Urocortin 3	Y	B40:01	0.3394
GQWEDASLLSK	IFN-α	0/4	4/4	11	Q969E3	46-56	UCN3	SPC,UCNIII	Urocortin 3	Y	A03:01	0.6854
SEYGFAEKV	IFN-α	0/4	4/4	9	Q6BDS2;A0JNW5	113-121	UHRF1BP1;UI	C6orf107,dJ349A	UHRF1 binding protein 1;UHRF1 binding protein	N	B49:01	0.0035
AEPKDFKRI	IFN-α	0/4	4/4	9	Q9UPW8	367-375	UNC13A	KIAA1032,Munc	Unc-13 homolog A	N	B49:01	0.2417
NLFHFVTDV	IFN-α	0/4	1/4	9	Q9UPW8	825-833	UNC13A	KIAA1032,Munc	Unc-13 homolog A	N	A02:01	0.1982
KELGQLSKL	IFN-α	0/4	2/4	9	Q9UPW8	1434-1442	UNC13A	KIAA1032,Munc	Unc-13 homolog A	N	B40:01/B49:01	0.0579/0.1047
IIGPQLSDK	Both	1/4	1/4	9	Q9UPW8	1573-1581	UNC13A	KIAA1032,Munc	Unc;13 homolog A	N	A03:01	0.6635
ATFPPFLHK	Both	2/4	4/4	9	Q9P2D8	183-191	UNC79	KIAA1409	Unc;79 homolog (C; elegans)	N	A03:01	0.0014
KEFPFPTTI	IFN-α	0/4	4/4	9	Q9P2D8	1185-1193	UNC79	KIAA1409	Unc-79 homolog (C. elegans)	N	B49:01	0.0014
SEFKIQIV	IFN-α	0/4	4/4	8	Q9P2D8	1827-1834	UNC79	KIAA1409	Unc-79 homolog (C. elegans)	N	B49:01	0.0309
REYQDLLNV	Both	4/4	4/4	9	P08670;Q16352	381-389	VIM;INA	NEF5,NF,66	Internexin neuronal intermediate filament prot	N	B49:01	0.0068
REYQDLLNVKM	IFN-α	0/4	4/4	11	P08670;Q16352	381-391	VIM;INA	NEF5,NF-66	Vimentin;Internexin neuronal intermediate filar	N	B40:01	0.1942
HLVLDLAQA	IFN-α	0/4	1/4	9	Q8N398	105-113	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	A02:01	0.8443
AARDSEQSTDALTD	Basal	1/4	0/4	14	Q8N398	624-637	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	NA	NA
SYIREQYVL	IFN-α	0/4	4/4	9	Q8N398	659-667	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	C07:01	0.5848
KVGALSTEVLG	Basal	1/4	0/4	11	Q8N398	734-744	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	NA	NA
ALSTEVLGR	Both	3/4	1/4	9	Q8N398	737-745	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	A03:01	0.7028
RQHRAALAGR	IFN-α	0/4	1/4	10	Q8N398	745-754	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	NA	NA
RAALAGRSL	Both	4/4	4/4	9	Q8N398	748-756	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	C03:04	0.0408
TAVPPSGEL	IFN-α	0/4	1/4	9	Q8N398	819-827	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	C03:04	0.0048
REVLPGALQV	IFN-α	0/4	1/4	10	Q8N398	947-956	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	B49:01	0.0475
VVYSKGLQR	Both	2/4	4/4	9	Q8N398	983-991	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	A03:01	0.0334
QEAPGSFRL	IFN-α	0/4	3/4	9	Q8N398	1064-1072	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	B40:01	0.0144
RASPFAVHR	Both	3/4	4/4	9	Q8N398	1089-1097	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	A03:01/C07:01	0.5132/1.0329
AEGLGGTDL	IFN-α	0/4	1/4	9	Q8N398	1154-1162	VWA5B2	DKFZp761K032,L	Von Willebrand factor A domain containing 5B2	N	B40:01	0.2003
SETDLERAV	IFN-α	0/4	1/4	9	O76024	168-176	WFS1	DFNA14,DFNA38	Wolframin ER transmembrane glycoprotein	N	B49:01	0.0339
SESKNYIAL	IFN-α	0/4	3/4	9	O76024	236-244	WFS1	DFNA14,DFNA38	Wolframin ER transmembrane glycoprotein	N	B40:01	0.0163
AEVNFGWNHL	IFN-α	0/4	2/4	10	O76024	393-402	WFS1	DFNA14,DFNA38	Wolframin ER transmembrane glycoprotein	N	B40:01	0.1323

Table S2, continued

ASFSVVGMVK	IFN-α	0/4	2/4	10	O76024	616-625	WFS1	DFNA14,DFNA3	Wolframin ER transmembrane glycoprotein	N	A03:01	0.1797
EEDDVTKDIVL	IFN-α	0/4	1/4	11	O76024	794-804	WFS1	DFNA14,DFNA3	Wolframin ER transmembrane glycoprotein	N	B40:01	0.4384
SEFKSVLL	IFN-α	0/4	4/4	8	O76024	808-815	WFS1	DFNA14,DFNA3	Wolframin ER transmembrane glycoprotein	N	B40:01/B49:01	0.1222/0.1244
IRSSKLIQH	IFN-α	0/4	1/4	9	O75123	329-337	ZNF623	KIAA0628	Zinc finger protein 623	N	C07:01	0.4547
RIHTGEKPYK	IFN-α	0/4	2/4	10	Q8TF20;Q9UEG4;	341-350	ZNF721;ZNF6	KIAA1982,KIAA0	Zinc finger protein 721;Zinc finger protein 629;Z	N	A03:01	0.1328

Table S3. Post-translationally modified HLA-I-eluted peptides identified in ECN90 β cells. The first column displays the peptide sequence with, in parenthesis, the mass shift of PTMs found on the aa, with PTMs detailed in the next column. Subsequent columns indicate the condition and the number of replicates in which the peptide were identified; aa length; UniProt accession number; aa position within the source protein; gene name, synonyms and description; whether the source is a granule protein; predicted HLA-I restrictions and ranking of unmodified sequences based on NetMHC4.1a; and the final PTM origin assigned as likely artifactual (i.e. found in both the modified and unmodified synthetic peptide) or likely biological (i.e. found only in the modified synthetic peptide). Y, yes; N, no; NA, not assigned; ND, not determined.

Peptide	PTM	Presence	Basal - # of replicates	IFN-α - # of replicates	Length	UniProt	Position	Gene name	Gene synonyms	Gene description	Granule protein	HLA restriction	Rank	Origin
KIYNKIM(+15.99)HL	Oxidation (M)	IFN-α	0/4	1/4	9	Q09428	394-402	ABCC8	ABC36,HHF1,HI,HRINS,I	ATP binding cassette subfamily C member 8	Y	A02:01	0.0403	Artifactual
AATSYIMC(+305.07)L	Glutathione disulfide	IFN-α	0/4	4/4	9	Q5TH69	1350-1358	ARFGEF3	BIG3,C6orf92,dJ171N1	ARFGEF family member 3	N	C03:04	0.2907	ND
KEYEPEM(+15.99)GKV	Oxidation (M)	IFN-α	0/4	4/4	10	Q93084;P166	120-129	ATP2A3;ATP2A2	SERCA3,ATP2B,DAR,SEI	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3;ATPa	N	B49:01	0.046	Artifactual
KEYEPEM(+15.99)GKVI	Oxidation (M)	IFN-α	0/4	2/4	11	Q93084	120-130	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3	N	B49:01	0.0809	Artifactual
YEPEM(+15.99)GKVI	Oxidation (M)	IFN-α	0/4	1/4	9	Q93084	122-130	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3	N	B49:01/B40:01	0.2008/0.3741	Artifactual
AEADAGSC(+305.07)LL	Glutathione disulfide	IFN-α	0/4	2/4	10	Q93084	370-379	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3	N	B40:01	0.0809	Biological
M(+15.99)RKEFTLEF	Oxidation (M)	IFN-α	0/4	3/4	9	Q93084	479-487	ATP2A3	SERCA3	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 3	N	C07:01	0.0045	Artifactual
QLM(+15.99)EKLREK	Oxidation (M)	IFN-α	0/4	1/4	9	Q9HBH7;Q9B	91-99	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-linked 2	N	A03:01	0.0744	Artifactual
Q(-17.03)LSHSLRAV	Pyro-glu from Q	IFN-α	0/4	3/4	9	Q9HBH7;Q9B	100-108	BEX1;BEX2	DJ79P11.1	Brain expressed X-linked 1;Brain expressed X-linked 2	N	A02:01	0.8772	Artifactual
QEVDGVM(+15.99)RI	Oxidation (M)	IFN-α	0/4	2/4	9	Q9NY47	79-87	CACNA2D2	KIAA0558	Calcium voltage-gated channel auxiliary subunit alpha2delta 2	N	B49:01	0.0051	Artifactual
SLFEM(+15.99)VQRL	Oxidation (M)	IFN-α	0/4	1/4	9	Q9ULU8	680-688	CADPS	CAPS,CAPS1,KIAA1121,I	Calcium dependent secretion activator	N	A02:01	0.0064	Artifactual
SLLERVLM(+15.99)K	Oxidation (M)	Both	3/4	3/4	9	Q9ULU8	815-823	CADPS	CAPS,CAPS1,KIAA1121,I	Calcium dependent secretion activator	N	A03:01	0.0095	Artifactual
LKEAVEEPSSKDVM(+15.99)	Oxidation (M)	Basal	1/4	0/4	14	P10645	117-130	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KEAVEEPSSKDVM(+15.99)E	Oxidation (M)	Basal	1/4	0/4	14	P10645	118-131	CHGA		Chromogranin A	Y	NA	NA	Artifactual
AVEEPSSKDVM(+15.99)	Oxidation (M)	Basal	1/4	0/4	11	P10645	120-130	CHGA		Chromogranin A	Y	B40:01	1.6169	Artifactual
AVEEPSSKDVM(+15.99)E	Oxidation (M)	Both	4/4	4/4	12	P10645	120-131	CHGA		Chromogranin A	Y	NA	NA	Artifactual
AVEEPSSKDVM(+31.99)E	Sulphone	Basal	1/4	0/4	12	P10645	120-131	CHGA		Chromogranin A	Y	NA	NA	ND
VEEPSSK(+14.02)DV	Methylation(KR)	IFN-α	0/4	1/4	9	P10645	121-129	CHGA		Chromogranin A	Y	B49:01	0.2372	Artifactual
VEEPSSKDVM(+15.99)	Oxidation (M)	Both	2/4	4/4	10	P10645	121-130	CHGA		Chromogranin A	Y	B40:01/B49:01	0.5364/1.5396	Artifactual
VEEPSSKDVM(+31.99)	Sulphone	IFN-α	0/4	1/4	10	P10645	121-130	CHGA		Chromogranin A	Y	B40:01/B49:01	0.5364/1.5396	ND
VEEPSSKDVM(+15.99)E	Oxidation (M)	Both	1/4	1/4	11	P10645	121-131	CHGA		Chromogranin A	Y	NA	NA	Artifactual
EEPSSKDVM(+15.99)E	Oxidation (M)	Basal	1/4	0/4	10	P10645	122-131	CHGA		Chromogranin A	Y	NA	NA	Artifactual
ALPEPMQ(+.98)ESK	Deamidation (NQ)	Both	1/4	1/4	10	P10645	153-162	CHGA		Chromogranin A	Y	A03:01	0.3446	Artifactual
ALPEPM(+15.99)QESK	Oxidation (M)	Both	3/4	3/4	10	P10645	153-162	CHGA		Chromogranin A	Y	A03:01	0.3446	Artifactual
ALPEPM(+15.99)Q(+.98)ESK	Oxidation (M); Deamidation (NQ)	Basal	1/4	0/4	10	P10645	153-162	CHGA		Chromogranin A	Y	A03:01	0.3446	Artifactual
L(+57.02)NPHPSLGY	Carbamidomethylation (DHKE X@N-term)	Basal	1/4	0/4	9	P10645	254-262	CHGA		Chromogranin A	Y	NA	NA	Artifactual
AVVPQ(+.98)GLFRGGK	Deamidation (NQ)	Basal	1/4	0/4	12	P10645	310-321	CHGA		Chromogranin A	Y	A03:01	0.355	Artifactual
S(+79.97)GELEQEEERLS	Phosphorylation (STY)	Basal	1/4	0/4	12	P10645	322-333	CHGA		Chromogranin A	Y	NA	NA	ND
Q(-17.03)EEERLSKEWEDS	Pyro-glu from Q	Basal	1/4	0/4	13	P10645	327-339	CHGA		Chromogranin A	Y	NA	NA	Artifactual
WSKM(+15.99)DQLAKE	Oxidation (M)	Basal	1/4	0/4	10	P10645	342-351	CHGA		Chromogranin A	Y	NA	NA	Artifactual
W(+31.99)SKMDQLAKELTAE	Dihydroxy	Both	1/4	1/4	14	P10645	342-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
WSKM(+15.99)DQLAKELTAE	Oxidation (M)	Basal	1/4	0/4	14	P10645	342-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
WSKM(+31.99)DQLAKELTAE	Sulphone	Basal	1/4	0/4	14	P10645	342-355	CHGA		Chromogranin A	Y	NA	NA	ND
SKM(+15.99)DQLAKE	Oxidation (M)	Basal	1/4	0/4	9	P10645	343-351	CHGA		Chromogranin A	Y	NA	NA	Artifactual
SKM(+15.99)DQLAKELT	Oxidation (M)	Basal	1/4	0/4	11	P10645	343-353	CHGA		Chromogranin A	Y	NA	NA	Artifactual
SKMDQ(+.98)LAKELTAE	Deamidation (NQ)	Basal	1/4	0/4	13	P10645	343-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
SKM(+15.99)DQLAKELTAE	Oxidation (M)	Both	2/4	1/4	13	P10645	343-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQLAKE	Oxidation (M)	Both	4/4	2/4	8	P10645	344-351	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQ(+.98)LAKEL	Deamidation (NQ); ;Oxidation (M)	Both	3/4	4/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
K(+42.01)MDQLAKEL	Acetylation (N-term)	IFN-α	0/4	2/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
K(+42.01)M(+15.99)DQLAKEL	Acetylation (N-term); Oxidation (M)	IFN-α	0/4	1/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
KMDQ(+.98)LAKEL	Deamidation (NQ)	Both	4/4	4/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
KMD(-18.01)QLAKEL	Dehydration	IFN-α	0/4	2/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
KM(+15.99)DQLAKEL	Oxidation (M)	Both	4/4	4/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	Artifactual
KM(+31.99)DQLAKEL	Sulphone	Both	3/4	1/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406	ND
KM(+15.99)DQLAKELT	Oxidation (M)	Both	3/4	4/4	10	P10645	344-353	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQ(+.98)LAKELTA	Deamidation (NQ); ;Oxidation (M)	Both	2/4	3/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA	Artifactual
K(+42.01)MDQLAKELTA	Acetylation (N-term)	Both	3/4	4/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KMDQ(+.98)LAKELTA	Deamidation (NQ)	Both	4/4	3/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQLAKELTA	Oxidation (M)	Both	4/4	4/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+31.99)DQLAKELTA	Sulphone	Both	2/4	1/4	11	P10645	344-354	CHGA		Chromogranin A	Y	NA	NA	ND
KMDQ(+.98)LAKELTAE	Deamidation (NQ)	Basal	1/4	0/4	12	P10645	344-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQLAKELTAE	Oxidation (M)	Both	4/4	4/4	12	P10645	344-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+15.99)DQ(+.98)LAKELTAE	Oxidation (M); Deamidation (NQ)	Basal	1/4	0/4	12	P10645	344-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
KM(+31.99)DQLAKELTAE	Sulphone	IFN-α	0/4	1/4	12	P10645	344-355	CHGA		Chromogranin A	Y	NA	NA	ND
M(+15.99)DQLAKEL	Oxidation (M)	Both	2/4	4/4	8	P10645	345-352	CHGA		Chromogranin A	Y	NA	NA	Artifactual
M(+15.99)DQLAKELT	Oxidation (M)	Basal	1/4	0/4	9	P10645	345-353	CHGA		Chromogranin A	Y	NA	NA	Artifactual
M(+15.99)DQLAKELTA	Oxidation (M)	Basal	1/4	0/4	10	P10645	345-354	CHGA		Chromogranin A	Y	NA	NA	Artifactual
M(+15.99)DQLAKELTAE	Oxidation (M)	Both	1/4	1/4	11	P10645	345-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
DQ(+.98)LAKELTAE	Deamidation (NQ)	Basal	1/4	0/4	10	P10645	346-355	CHGA		Chromogranin A	Y	NA	NA	Artifactual
Q(-17.03)EEEEDNRDSSMKL	Pyro-glu from Q	Both	3/4	3/4	14	P10645	361-374	CHGA		Chromogranin A	Y	NA	NA	Artifactual
EEEEDNRDSSM(+15.99)KL	Oxidation (M)	Basal	1/4	0/4	13	P10645	362-374	CHGA		Chromogranin A	Y	NA	NA	Artifactual
EEDNRDSSM(+15.99)KL	Oxidation (M)	Basal	1/4	0/4	11	P10645	364-374	CHGA		Chromogranin A	Y	B40:01/B49:01	0.6453/1.7105	Artifactual
AR(+28.03)AYGFRGPGPQL	Dimethylation(KR)	Both	1/4	1/4	13	P10645	378-390	CHGA		Chromogranin A	Y	NA	NA	Artifactual
RAYGFRGPGPQ(+.98)	Deamidation (NQ)	Basal	3/4	0/4	11	P10645	379-389	CHGA		Chromogranin A	Y	NA	NA	Artifactual
RAYGFRGPGPQ(-18.01)	Dehydration	Basal	1/4	0/4	11	P10645	379-389	CHGA		Chromogranin A	Y	NA	NA	Artifactual
RAYGFRGPGPQ(+.98)L	Deamidation (NQ)	Basal	2/4	0/4	12	P10645	379-390	CHGA		Chromogranin A	Y	C03:04	0.8606	Artifactual
RAYGFRGPGPQ(+.98)LR	Deamidation (NQ)	Both	4/4	2/4	13	P10645	379-391	CHGA		Chromogranin A	Y	A03:01	1.7153	Artifactual
Y(+42.01)GFRGPGPQL	Acetylation (N-term)	Basal	1/4	0/4	10	P10645	381-390	CHGA		Chromogranin A	Y	C03:04	0.4651	Artifactual
Y(+183.04)GFRGPGPQL	Aminoethylbenzenesulfonylation	Both	4/4	4/4	10	P10645	381-390	CHGA		Chromogranin A	Y	C03:04	0.4651	Artifactual
YGFRGPGPQ(+.98)L	Deamidation (NQ)	Both	4/4	4/4	10	P10645	381-390	CHGA		Chromogranin A	Y	C03:04	0.4651	Artifactual
GFRGPGPQ(+.98)L	Deamidation (NQ)	Both	3/4	1/4	9	P10645	382-390	CHGA		Chromogranin A	Y	C07:01	1.3265	Artifactual
GFRGPGPQ(+.98)LR	Deamidation (NQ)	Basal	2/4	0/4	10	P10645	382-391	CHGA		Chromogranin A	Y	NA	NA	Artifactual
FRGPGPQ(+.98)L	Deamidation (NQ)	Both	4/4	4/4	8	P10645	383-390	CHGA		Chromogranin A	Y	C07:01	0.1921	Artifactual
SSREDSLEAGLPLQ(+.98)	Deamidation (NQ)	Basal	1/4	0/4	14	P10645	397-410	CHGA		Chromogranin A	Y	NA	NA	Artifactual
SSREDS(+79.97)LEAGLPLQ	Phosphorylation (STY)	Basal	1/4	0/4	14	P10645	397-410	CHGA		Chromogranin A	Y	NA	NA	ND
SREDSLEAGLPLQ(+.98)V	Deamidation (NQ)	Basal	1/4	0/4	14	P10645	398-411	CHGA		Chromogranin A	Y	NA	NA	Artifactual
S(+79.97)REDSLEAGLPLQV	Phosphorylation (STY)	Basal	1/4	0/4	14	P10645	398-411	CHGA		Chromogranin A	Y	NA	NA	ND
REDSLEAGLPLQ(+.98)V	Deamidation (NQ)	Basal	1/4	0/4	13	P10645	399-411	CHGA		Chromogranin A	Y	B49:01	1.4346	Artifactual
EDS(+79.97)LEAGLPLQV	Phosphorylation (STY)	Basal	1/4	0/4	12	P10645	400-411	CHGA		Chromogranin A	Y	NA	NA	ND
EDS(+79.97)LEAGLPLQVR	Phosphorylation (STY)	Basal	1/4	0/4	13	P10645	400-412	CHGA		Chromogranin A	Y	NA	NA	ND
LEAGLPLQ(+.98)V	Deamidation (NQ)	Both	1/4	2/4	9	P10645	403-411	CHGA		Chromogranin A	Y	B49:01	0.0173	Artifactual
RVAQ(+.98)LDQLLHY	Deamidation (NQ)	Basal	1/4	0/4	11	P05060	603-613	CHGB	SCG1,Sgl	Chromogranin B	Y	A03:01	0.6234	Artifactual
RVAQLDQ(+.98)LLHY	Deamidation (NQ)	Basal	1/4	0/4	11	P05060	603-613	CHGB	SCG1,Sgl	Chromogranin B	Y	A03:01	0.6234	Artifactual
AQ(+.98)LDQLLHY	Deamidation (NQ)	Basal	3/4	0/4	9	P05060	605-613	CHGB	SCG1,Sgl	Chromogranin B	Y	C07:01/A03:01	0.309/0.715	Artifactual

Table S3, continued

AQLDQ(+.98)LLHY	Deamidation (NQ)	Basal	1/4	0/4	9	P05060	605-613	CHGB	SCG1,Sgl	Chromogranin B	Y	C07:01/A03:01	0.309/0.715	Artifactual
VLYRPDGQ(+.98)HDGK	Deamidation (NQ)	Basal	1/4	0/4	12	Q12860	942-953	CNTN1	GP135	Contactin 1	N	A03:01	1.2858	Artifactual
KM(+15.99)MSETLNF	Oxidation (M)	Basal	1/4	0/4	9	P16870	468-476	CPE		Carboxypeptidase E	Y	E01:01	0.1764	Artifactual
KM(+31.99)MSETLNF	Sulphone	Basal	3/4	0/4	9	P16870	468-476	CPE		Carboxypeptidase E	Y	E01:01	0.1764	ND
KMM(+31.99)SETLNF	Sulphone	IFN-α	0/4	1/4	9	P16870	468-476	CPE		Carboxypeptidase E	Y	E01:01	0.1764	ND
FADDPC(+305.07)SSV	Glutathione disulfide	IFN-α	0/4	4/4	9	P35221;P262	111-119	CTNNA1;CTNNA	CAP102,CAP-R,CT114	Catenin alpha 1;Catenin alpha 2	N	C03:04/A02:01	0.1577/0.4548	Biological
KM(+15.99)KAPEKKPLVK	Oxidation (M)	Basal	4/4	0/4	12	P35221;P262	860-871	CTNNA1;CTNNA	CAP102,CAP-R,CT114	Catenin alpha 1;Catenin alpha 2	N	A03:01	0.272	Artifactual
AIAEAM(+15.99)KVK	Oxidation (M)	Both	4/4	4/4	9	Q9UI36	337-345	DACH1	DACH	Dachshund family transcription factor 1	N	A03:01	0.1576	Artifactual
SLRRSGRQ(+.98)PK	Deamidation (NQ)	Basal	1/4	0/4	10	Q9BTC0	60-69	DIDO1	BYE1,C20orf158,DATF1	Death inducer-obliterator 1	N	A03:01	0.4291	Artifactual
M(+15.99)EDGLIKQI	Oxidation (M)	IFN-α	0/4	1/4	9	Q16555	37-45	DPYSL2	CRMP2,DHPRP2,DRP-2,	Dihydropyrimidinase like 2	N	B49:01	0.0126	Artifactual
GLYDGPVC(+119.00)EV	Cysteinylation	IFN-α	0/4	1/4	10	Q16555	497-506	DPYSL2	CRMP2,DHPRP2,DRP-2,	Dihydropyrimidinase like 2	N	A02:01	0.0129	Biological
SLSGAQ(+.98)IDDNIPR	Deamidation (NQ)	Basal	1/4	0/4	13	Q16555	540-552	DPYSL2	CRMP2,DHPRP2,DRP-2,	Dihydropyrimidinase like 2	N	NA	NA	Artifactual
KM(+15.99)FPEVKEK	Oxidation (M)	IFN-α	0/4	4/4	9	Q05329	257-265	GAD2	GAD65	Glutamate decarboxylase 2	Y	A03:01	0.0069	Artifactual
KM(+15.99)FDVGGQR	Oxidation (M)	Basal	1/4	0/4	9	P04899;P087	198-206	GNAI2;GNAI3;GN	GIP,GNAI2B,87U6,	G protein subunit alpha i2;G protein subunit alpha i3;G protein subunit	N	A03:01	0.1598	Artifactual
GESGK(+42.01)STIV	Acetylation (K)	IFN-α	0/4	1/4	9	Q5JWF2;P630	692-700	GNAS;GNAS;GNA	GNAS1,GNASXL,GPSA,N	GNAS complex locus;GNAS complex locus;G protein subunit alpha i2;G	N	B49:01/B40:01	0.1615/0.308	Artifactual
SEYPIVD(-18.01)GKL	Dehydration	IFN-α	0/4	4/4	10	P54868;Q015	248-257	HMGCS2;HMGCS	HMGCS	3-hydroxy-3-methylglutaryl-CoA synthase 1;3-hydroxy-3-methylglutary	N	B40:01/B49:01	0.0853/0.2371	Artifactual
KLSLVAAM(+15.99)L	Oxidation (M)	IFN-α	0/4	1/4	9	P11021	2-10	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A02:01	0.8236	Artifactual
KLSLVAAM(+31.99)L	Sulphone	IFN-α	0/4	1/4	9	P11021	2-10	HSPA5	BiP,GRP78	Heat shock protein family A (Hsp70) member 5	N	A02:01	0.8236	Biological
GIPPAPR(+14.02)GV	Methylation(KR)	Basal	3/4	0/4	9	P11021;P111	486-494	HSPA5;HSPA8;HS	HSC70,HSC71,HSP73,HS	Heat shock protein family A (Hsp70) member 8;Heat shock protein fam	N	A02:01	1.5107	Artifactual
A(+42.01)LWGPDPAAA	Acetylation (Protein N-term)	IFN-α	0/4	4/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALW(+43.99)GPDPAAA	Carboxylation (DKW)	Both	1/4	1/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALW(+31.99)GPDPAAA	Dihydroxy	Both	2/4	3/4	10	P01308;F8WC	15-24	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.3268	Artifactual
A(+27.99)LWGPDPAAA	Formylation (Protein N-term)	IFN-α	0/4	1/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALW(+15.99)GPDPAAA	Oxidation (HW)	Both	4/4	4/4	10	P01308;F8WC	15-24	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.3268	Artifactual
ALWGPD(+37.95)PAAA	Replacement of 2 protons by calcium	Both	3/4	1/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALWGPD(+21.97)PAAA	Replacement of 2 protons by magnesium	Both	4/4	4/4	10	P01308;F8WC	15-24	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.3268	Artifactual
ALWGPDPAAA(+17.03)	Replacement of proton with ammonium ion	Both	1/4	2/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALWGPD(+21.98)PAAA	Sodium adduct	Both	2/4	2/4	10	P01308;F8WC	15-24	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	A02:01	0.3268	Artifactual
ALW(+3.99)GPDPAAA	Tryptophan oxidation to kynurenin	Both	4/4	4/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
ALW(+13.98)GPDPAAA	Tryptophan oxidation to oxolactone	Both	1/4	1/4	10	P01308	15-24	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.3268	Artifactual
L(+57.02)WGPDPAAA	Carbamidomethylation (DHKE X@N-term)	Basal	1/4	0/4	9	P01308;F8WC	16-24	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	Artifactual
FVNQHLC(+70.04)GS	Crotonaldehyde	Basal	3/4	0/4	9	P01308;F8WC	25-33	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	Artifactual
FVNQHLC(+119.00)GS	Cysteinylation	Basal	1/4	0/4	9	P01308;F8WC	25-33	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
FVNQHLC(+305.07)GS	Glutathione disulfide	Basal	1/4	0/4	9	P01308	25-33	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
FVNQHLC(+70.04)GSH	Crotonaldehyde	Both	1/4	3/4	10	P01308;F8WC	25-34	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	Artifactual
FVNQHLC(+47.98)GSH	Cysteine oxidation to cysteic acid	Basal	2/4	0/4	10	P01308	25-34	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
FVNQHLC(+119.00)GSH	Cysteinylation	Basal	1/4	0/4	10	P01308	25-34	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
FVNQHLC(+305.07)GSH	Glutathione disulfide	Basal	2/4	0/4	10	P01308	25-34	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
FVNQHLC(+70.04)GSHL	Crotonaldehyde	IFN-α	0/4	1/4	11	P01308	25-35	INS	IDDM1,IDDM2	Insulin	Y	C03:04	1.9425	Artifactual
FVNQHLC(+305.07)GSHL	Glutathione disulfide	Both	3/4	3/4	11	P01308	25-35	INS	IDDM1,IDDM2	Insulin	Y	C03:04	1.9425	ND
FVNQHLC(+305.07)GSHLVE	Glutathione disulfide	Basal	1/4	0/4	13	P01308;F8WC	25-37	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
FVNQHLC(+70.04)GSHLVEA	Crotonaldehyde	Both	2/4	2/4	14	P01308	25-38	INS	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	Artifactual
FVNQHLC(+47.98)GSHLVEA	Cysteine oxidation to cysteic acid	IFN-α	0/4	1/4	14	P01308	25-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
FVNQHLC(+119.00)GSHLVEA	Cysteinylation	Basal	2/4	0/4	14	P01308	25-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
FVNQHLC(+305.07)GSHLVEA	Glutathione disulfide	Both	3/4	2/4	14	P01308	25-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
VNQHLC(+305.07)GSHLVEA	Glutathione disulfide	Basal	1/4	0/4	13	P01308;F8WC	26-38	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
HLC(+70.04)GSHLVEA	Crotonaldehyde	Both	4/4	4/4	10	P01308;F8WC	29-38	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	Artifactual
HLC(+47.98)GSHLVEA	Cysteine oxidation to cysteic acid	IFN-α	0/4	2/4	10	P01308	29-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
HLC(+119.00)GSHLVEA	Cysteinylation	Both	1/4	3/4	10	P01308	29-38	INS	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
HLC(+31.99)GSHLVEA	Dihydroxy	IFN-α	0/4	3/4	10	P01308	29-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
HLC(+305.07)GSHLVEA	Glutathione disulfide	Both	1/4	2/4	10	P01308;F8WC	29-38	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
HLC(+305.07)GSHLVEALY	Glutathione disulfide	Basal	1/4	0/4	12	P01308;F8WC	29-40	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
HLC(+47.98)GSHLVEALYL	Cysteine oxidation to cysteic acid	IFN-α	0/4	1/4	13	P01308	29-41	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
LC(+119.00)GSHLVEA	Cysteinylation	Both	1/4	1/4	9	P01308	30-38	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
LC(+305.07)GSHLVEA	Glutathione disulfide	Basal	1/4	0/4	9	P01308;F8WC	30-38	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
HLVEALY(-18.01)L	Dehydration	Both	3/4	4/4	8	P01308	34-41	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
HLVEALYLV(-.98)	Amidation	IFN-α	0/4	1/4	9	P01308	34-42	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.0364	Artifactual
HLVEALY(-18.01)LV	Dehydration	IFN-α	0/4	1/4	9	P01308	34-42	INS	IDDM1,IDDM2	Insulin	Y	A02:01	0.0364	Artifactual
HLVEALYLVC(+119.00)	Cysteinylation	Basal	1/4	0/4	10	P01308	34-43	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
HLVEALYLVC(+47.98)GERG	Cysteine oxidation to cysteic acid	IFN-α	0/4	1/4	14	P01308	34-47	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
HLVEALYLVC(+305.07)GERG	Glutathione disulfide	Both	4/4	4/4	14	P01308	34-47	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
LVEALYLVC(+305.07)GERG	Glutathione disulfide	Both	1/4	1/4	13	P01308	35-47	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
LVEALYLVC(+305.07)GERGF	Glutathione disulfide	Basal	1/4	0/4	14	P01308	35-48	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
VEALYLVC(+305.07)GERG	Glutathione disulfide	Basal	1/4	0/4	12	P01308;F8WC	36-47	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
ALYLVC(+305.07)GERG	Glutathione disulfide	Both	2/4	1/4	10	P01308;F8WC	38-47	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
ALYLVC(+305.07)GERGF	Glutathione disulfide	Both	3/4	2/4	11	P01308	38-48	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
ALYLVC(+305.07)GERGFF	Glutathione disulfide	Both	3/4	1/4	12	P01308;F8WC	38-49	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
ALYLVC(+305.07)GERGFFY	Glutathione disulfide	Basal	1/4	0/4	13	P01308	38-50	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
LYLVC(+305.07)GERG	Glutathione disulfide	Basal	1/4	0/4	9	P01308;F8WC	39-47	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
YLVC(+305.07)GERGFFYTPK	Glutathione disulfide	IFN-α	0/4	1/4	14	P01308;F8WC	40-53	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
LVC(+305.07)GERGF	Glutathione disulfide	Basal	1/4	0/4	8	P01308;F8WC	41-48	INS;INS-IGF2	IDDM1,IDDM2,	Insulin;INS-IGF2 readthrough	Y	NA	NA	ND
LVC(+47.98)GERGFFYTPK	Cysteine oxidation to cysteic acid	Both	3/4	1/4	13	P01308	41-53	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
LVC(+305.07)GERGFFYTPK	Glutathione disulfide	Both	2/4	4/4	13	P01308;F8WC	41-53	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
LVC(+305.07)GERGFFYTPKT	Glutathione disulfide	Both	2/4	1/4	14	P01308;F8WC	41-54	INS;INS;IGF2	IDDM1,IDDM2	Insulin;INS;IGF2 readthrough	Y	NA	NA	ND
VC(+47.98)GERGFFYTPK	Cysteine oxidation to cysteic acid	Basal	1/4	0/4	12	P01308	42-53	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
VC(+70.04)GERGFFYTPKT	Crotonaldehyde	Both	1/4	1/4	13	P01308	42-54	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
C(+12.00)GERGFFYTPKT	Addition of Carbon to cysteine	Basal	1/4	0/4	12	P01308	43-54	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
REAEDLQVGQ(+.98)V	Deamidation (NQ)	IFN-α	0/4	2/4	11	P01308	56-66	INS	IDDM1,IDDM2	Insulin	Y	B49:01/B40:01	0.1983/0.5284	Artifactual
REAEDLQVGQV(+14.02)	Methylation(C-term)	IFN-α	0/4	2/4	11	P01308	56-66	INS	IDDM1,IDDM2	Insulin	Y	B49:01/B40:01	0.1983/0.5284	Artifactual
REAEDLQ(+.98)VGQVEL	Deamidation (NQ)	IFN-α	0/4	3/4	13	P01308	56-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.2255	Artifactual
REAEDLQVGQ(+.98)VEL	Deamidation (NQ)	IFN-α	0/4	4/4	13	P01308	56-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.2255	Artifactual
AEDLQ(+.98)VGQVEL	Deamidation (NQ)	IFN-α	0/4	2/4	11	P01308	58-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.1346	Artifactual
AEDLQVGQ(+.98)VEL	Deamidation (NQ)	IFN-α	0/4	4/4	11	P01308	58-68	INS	IDDM1,IDDM2	Insulin	Y	B40:01	0.1346	Artifactual
LQ(+.98)VGQVEL	Deamidation (NQ)	IFN-α	0/4	1/4	8	P01308	61-68	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
LQVGQ(+.98)VEL	Deamidation (NQ)	IFN-α	0/4	1/4	8	P01308	61-68	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
GIVEQCCT(-2.02)S	2-amino-3-oxo-butanoic_acid	IFN-α	0/4	1/4	9	P01308	90-98	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual

Table S3, continued

GIVEQCCT(+117.00)SICSL	Phospho-propargylamine	Basal	2/4	0/4	13	P01308	90-102	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	Artifactual
SLYQLENYC(+305.07)N	Glutathione disulfide	Both	4/4	3/4	10	P01308	101-110	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
LYQLENYC(+305.07)N	Glutathione disulfide	Basal	1/4	0/4	9	P01308	102-110	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
YQLENYC(+305.07)N	Glutathione disulfide	Both	4/4	4/4	8	P01308	103-110	INS	IDDM1,IDDM2	Insulin	Y	NA	NA	ND
AVFN(+.98)IIFTQK	Deamidation (NQ)	IFN-α	0/4	4/4	10	Q12756	219-228	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	A03:01	0.022	Artifactual
AVFNIIFTQ(+.98)K	Deamidation (NQ)	IFN-α	0/4	4/4	10	Q12756	219-228	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	A03:01	0.022	Artifactual
L(+57.02)YPVPLVH	Carbamidomethylation (DHKE X@N-term)	IFN-α	0/4	1/4	8	Q12756	861-868	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	NA	NA	Artifactual
G(+42.01)EVKGFLRV	Acetylation (N-term)	IFN-α	0/4	1/4	9	Q12756	877-885	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	B49:01	0.0161	Artifactual
SESC(+305.07)PVVGM	Glutathione disulfide	IFN-α	0/4	4/4	9	Q12756	923-931	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	B40:01/B49:01	0.212/0.431	Biological
KLSEM(+15.99)SVTL	Oxidation (M)	Both	1/4	1/4	9	Q12756	1480-1488	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	A02:01	0.0117	Artifactual
KLSEM(+31.99)SVTL	Sulphone	IFN-α	0/4	1/4	9	Q12756	1480-1488	KIF1A	ATSV,C2orf20,SPG30,U	Kinesin family member 1A	N	A02:01	0.0117	Artifactual
S(+42.01)TNENANTPAARL	Acetylation (Protein N-term)	Basal	1/4	0/4	13	P52292	2-14	KPNA2	IPOA1,QIP2,RCH1,SRP1	Karyopherin subunit alpha 2	N	NA	NA	Artifactual
S(+42.01)TNENANTPAARLH	Acetylation (Protein N-term)	Basal	1/4	0/4	14	P52292	2-15	KPNA2	IPOA1,QIP2,RCH1,SRP1	Karyopherin subunit alpha 2	N	NA	NA	Artifactual
T(+57.02)NPKTNIQK	Carbamidomethylation (DHKE X@N-term)	IFN-α	0/4	1/4	9	P52292	345-353	KPNA2	IPOA1,QIP2,RCH1,SRP1	Karyopherin subunit alpha 2	N	NA	NA	Artifactual
SM(+15.99)KPEPLFR	Oxidation (M)	Both	2/4	2/4	9	P46821	392-400	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	A03:01	0.2094	Artifactual
FRASVLFDTM(+15.99)	Oxidation (M)	IFN-α	0/4	1/4	10	P46821	1832-1841	MAP1B	MAP5,PPP1R102	Microtubule associated protein 1B	N	C07:01	0.7701	Artifactual
AET(-18.01)IFVKV	Dehydration	IFN-α	0/4	2/4	8	O76050	335-342	NEURL1	h-neu,neu-1,NEURL,RN	Neuralized E3 ubiquitin protein ligase 1	N	B49:01	0.0924	Artifactual
KTDPKM(+15.99)LSK	Oxidation (M)	Both	2/4	3/4	9	Q9UBD5	194-202	ORC3	IMAGE50150,LATHEO,(	Origin recognition complex subunit 3	N	A03:01	0.0074	Artifactual
Q(-17.03)IGSLENHYLFK	Pyro-glu from Q	Basal	1/4	0/4	12	P29120	60-71	PCSK1	NEC1,PC1,PC3,SPC3	Proprotein convertase subtilisin/kexin type 1	Y	NA	NA	Artifactual
SALRDSALNLFND(-.98)	Amidation	Basal	1/4	0/4	13	P29120	111-123	PCSK1	NEC1,PC1,PC3,SPC3	Proprotein convertase subtilisin/kexin type 1	Y	NA	NA	Artifactual
GLFRPPPALC(+305.07)	Glutathione disulfide	Basal	1/4	0/4	10	Q9UHG2	24-33	PCSK1N	BigLEN,PEN,proSAAS,SA	Proprotein convertase subtilisin/kexin type 1 inhibitor	Y	A02:01	1.0864	Biological
ALIGQC(+305.07)IHI	Glutathione disulfide	IFN-α	0/4	1/4	9	Q86TG7	229-237	PEG10	HB-1,KIAA1051,Mar2,N	Paternally expressed 10	N	A02:01	0.2398	Biological
RVRQS(+79.97)PLATR	Phosphorylation (STY)	IFN-α	0/4	3/4	10	O75381	40-49	PEX14		Peroxisomal biogenesis factor 14	N	A03:01	0.1376	Biological
GLFGQC(+305.07)QVGV	Glutathione disulfide	IFN-α	0/4	1/4	10	Q16849	57-66	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	A02:01	0.2086	Biological
WESGC(+305.07)TVIV	Glutathione disulfide	IFN-α	0/4	2/4	9	Q16849	799-807	PTPRN	IA-2	Protein tyrosine phosphatase, receptor type N	Y	B49:01/B40:01	0.6306/0.765	Artifactual
KTITQ(+.98)IVKDK	Deamidation (NQ)	IFN-α	0/4	1/4	10	Q8HWS3	108-117	RFX6	dJ955L16.1,MGC33442	Regulatory factor X6	N	A03:01	0.0871	Artifactual
ATFGKTIRQ(+.98)K	Deamidation (NQ)	Both	2/4	3/4	10	Q8HWS3	164-173	RFX6	MGC33442,RFXDC1,dJ9	Regulatory factor X6	N	A03:01	0.0066	Artifactual
SM(+15.99)VSDIERV	Oxidation (M)	Both	3/4	4/4	9	Q8HWS3	408-416	RFX6	MGC33442,RFXDC1,dJ9	Regulatory factor X6	N	A02:01	0.0681	Artifactual
SM(+31.99)VSDIERV	Sulphone	IFN-α	0/4	1/4	9	Q8HWS3	408-416	RFX6	dJ955L16.1,MGC33442	Regulatory factor X6	N	A02:01	0.0681	Biological
KIAGYVTHLM(+15.99)K	Oxidation (M)	IFN-α	0/4	1/4	11	P08708	49-59	RPS17	MGC72007,RPS17L,RPS	Ribosomal protein S17	N	A03:01	0.0803	Artifactual
IEVDPDTKEM(+15.99)L	Oxidation (M)	IFN-α	0/4	4/4	11	P08708	96-106	RPS17	MGC72007,RPS17L,RPS	Ribosomal protein S17	N	B40:01	0.1571	Artifactual
IEVDPDTKEM(+31.99)L	Sulphone	IFN-α	0/4	1/4	11	P08708	96-106	RPS17	MGC72007,RPS17L,RPS	Ribosomal protein S17	N	B40:01	0.1571	Artifactual
NLQVTQPTVGM(+15.99)NF	Oxidation (M)	IFN-α	0/4	1/4	13	P08708	116-128	RPS17	MGC72007,RPS17L,RPS	Ribosomal protein S17	N	NA	NA	Artifactual
Q(-17.03)EKVLPRLPY	Pyro-glu from Q	Basal	1/4	0/4	10	P13521	459-468	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA	Artifactual
M(+15.99)LVKYPEIINSNQV	Oxidation (M)	Basal	1/4	0/4	14	P13521	511-524	SCG2	CHGC,SgII,SN	Secretogranin II	Y	NA	NA	Artifactual
VEYPAHQAM(+15.99)	Oxidation (M)	IFN-α	0/4	3/4	9	P05408	58-66	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B40:01/B49:01	0.0341/0.0761	Artifactual
HLFDPEHDY(+79.96)PG	Sulfation	Basal	1/4	0/4	11	P05408	149-159	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
LFDPEHDY(+79.96)PGLGKW	Sulfation	Basal	1/4	0/4	14	P05408	150-163	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
HDY(+79.96)PGLGKW	Sulfation	Basal	1/4	0/4	9	P05408	155-163	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	B49:01	1.3213	Artifactual
N(+.98)KKLLYEKMKGGE	Deamidation (NQ)	Basal	1/4	0/4	13	P05408	164-176	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
N(+.98)KKLLYEKM(+15.99)KGGE	Deamidation (NQ); Oxidation (M)	Basal	1/4	0/4	13	P05408	164-176	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
NKKLLYEKM(+15.99)KGGE	Oxidation (M)	Basal	1/4	0/4	13	P05408	164-176	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
SVNPYLQ(+.98)GQR	Deamidation (NQ)	Both	1/4	1/4	10	P05408	182-191	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.4187	Artifactual
SVNPYLQGQ(+.98)R	Deamidation (NQ)	Basal	3/4	0/4	10	P05408	182-191	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.4187	Artifactual
SVNPYLQ(+.98)GQRL	Deamidation (NQ)	Basal	3/4	0/4	11	P05408	182-192	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
SVNPYLQGQ(+.98)RL	Deamidation (NQ)	Basal	1/4	0/4	11	P05408	182-192	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
SVNPYLQ(+.98)GQRLDN	Deamidation (NQ)	Basal	2/4	0/4	13	P05408	182-194	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
SVNPYLQGQ(+.98)RLDN	Deamidation (NQ)	Basal	2/4	0/4	13	P05408	182-194	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
SVNPYLQGQRLDN(+.98)	Deamidation (NQ)	Basal	2/4	0/4	13	P05408	182-194	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	Artifactual
VVAKK(+42.01)SVPH	Acetylation (K)	IFN-α	0/4	1/4	9	P05408	195-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.0522	Artifactual
V(+42.01)VAKKSVPH	Acetylation (N-term)	Both	2/4	4/4	9	P05408	195-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.0522	Artifactual
VVAKKS(+42.01)VPH	Acetylation (TSCYH)	Both	3/4	3/4	9	P05408	195-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.0522	Artifactual
VVAKKSVPH(+70.04)	Crotonaldehyde	Both	3/4	3/4	9	P05408	195-203	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	A03:01	1.0522	Artifactual
SVPHFS(+79.97)DEDKDPE	Phosphorylation (STY)	Basal	1/4	0/4	13	P05408	200-212	SCG5	7B2,SGNE1,SgV	Secretogranin V	Y	NA	NA	ND
KAFQ(+.98)ATQQK	Deamidation (NQ)	IFN-α	0/4	1/4	9	Q8WTS6	358-366	SETD7	KIAA1717,KMT7,SET7,S	SET domain containing lysine methyltransferase 7	N	A03:01	0.0732	Artifactual
KAFQATQ(+.98)QK	Deamidation (NQ)	IFN-α	0/4	4/4	9	Q8WTS6	358-366	SETD7	KIAA1717,KMT7,SET7,S	SET domain containing lysine methyltransferase 7	N	A03:01	0.0732	Artifactual
RTYLVNDKAAKM(+15.99)	Oxidation (M)	Basal	2/4	0/4	12	Q8IWU4	6-17	SLC30A8	ZnT-8,ZNT8	Solute carrier family 30 member 8	Y	NA	NA	Artifactual
SLWSEPVVC(+119.00)	Cysteinylation	IFN-α	0/4	1/4	9	Q9UHI5	441-449	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	A02:01	0.3791	ND
SLWSEPVVC(+305.07)	Glutathione disulfide	IFN-α	0/4	4/4	9	Q9UHI5	441-449	SLC7A8	LAT2,LPI-PC1	Solute carrier family 7 member 8	N	A02:01	0.3791	Biological
KRAS(+79.97)GQAFEL	Phosphorylation (STY)	IFN-α	0/4	1/4	10	Q93045;P169	47-56	STMN2;STMN1	C1orf215,FLJ32206,Lag	Stathmin 1;Stathmin 2	N	C07:01	0.5721	ND
KIM(+15.99)HDVIKK	Oxidation (M)	Both	3/4	2/4	9	P61764	13-21	STXBP1	MUNC18,UNC18,hUNC	Syntaxin binding protein 1	N	A03:01	0.0154	Artifactual
KLKIEQM(+15.99)QK	Oxidation (M)	IFN-α	0/4	1/4	9	Q9BTF0	146-154	THUMPD2	C2orf8,MGC2454	THUMP domain containing 2	N	A03:01	0.0359	Artifactual
Q(+.98)EVGKVIGI	Deamidation (NQ)	IFN-α	0/4	3/4	9	Q9BTF0	224-232	THUMPD2	C2orf8,MGC2454	THUMP domain containing 2	N	B49:01	0.0419	Artifactual
KVIGIAIM(+15.99)K	Oxidation (M)	Both	3/4	4/4	9	Q9BTF0	228-236	THUMPD2	C2orf8,MGC2454	THUMP domain containing 2	N	A03:01	0.0117	Artifactual
KVIGIAIM(+31.99)K	Sulphone	IFN-α	0/4	1/4	9	Q9BTF0	228-236	THUMPD2	C2orf8,MGC2454	THUMP domain containing 2	N	A03:01	0.0117	Artifactual
KEW(+15.99)PDVYYV	Oxidation (HW)	IFN-α	0/4	1/4	9	Q9BTF0	318-326	THUMPD2	C2orf8,MGC2454	THUMP domain containing 2	N	B49:01	0.0292	Artifactual
AELC(+305.07)DIAAI	Glutathione disulfide	IFN-α	0/4	2/4	9	P28289	126-134	TMOD1	D9S57E,ETMOD,TMOD	Tropomodulin 1	N	B49:01/B40:01	0.1865/0.4569	Biological
ALSSSSIM(+15.99)NK	Oxidation (M)	IFN-α	0/4	2/4	10	P28289	149-158	TMOD1	D9S57E,ETMOD,TMOD	Tropomodulin 1	N	A03:01	0.0298	Artifactual
STTAVVTN(+.98)PK	Deamidation (NQ)	Basal	2/4	0/4	10	P02766	137-146	TTR	CTS,CTS1,HsT2651,PALI	Transthyretin	N	A03:01	0.577	Artifactual
M(+31.99)LMPVHFLL	Sulphone	IFN-α	0/4	1/4	9	Q969E3	9-21	UCN3	SPC,UCNIII	Urocortin 3	Y	A02:01	0.0379	ND
GQ(+.98)WEDASLLSK	Deamidation (NQ)	IFN-α	0/4	1/4	11	Q969E3	46-56	UCN3	SPC,UCNIII	Urocortin 3	Y	A03:01	0.6854	Artifactual
REYQ(+.98)DLLNV	Deamidation (NQ)	IFN-α	0/4	1/4	9	P08670;Q163	381-389	VIM;INA	NEF5,NF-66	Vimentin;Internexin neuronal intermediate filament protein alpha	N	B49:01	0.0068	Artifactual
REYQDLLNVKM(+15.99)	Oxidation (M)	IFN-α	0/4	2/4	11	P08670;Q163	381-391	VIM;INA	NEF5,NF-66	Vimentin;Internexin neuronal intermediate filament protein alpha	N	B40:01	0.1942	Artifactual
SC(+305.07)PLPAPTPAPF	Glutathione disulfide	Basal	1/4	0/4	12	Q8N398	722-733	VWA5B2	DKFZp761K032,LOC901	Von Willebrand factor A domain containing 5B2	N	NA	NA	ND
HLLC(+305.07)YSPANV	Glutathione disulfide	Basal	1/4	0/4	10	Q8N398	1233-1242	VWA5B2	DKFZp761K032,LOC901	Von Willebrand factor A domain containing 5B2	N	NA	NA	ND
ASFSVVGM(+15.99)VK	Oxidation (M)	IFN-α	0/4	1/4	10	O76024	616-625	WFS1	DFNA14,DFNA38,DFNA	Wolframin ER transmembrane glycoprotein	N	A03:01	0.1797	Artifactual

Table S4. *Cis*-spliced HLA-I-eluted peptides identified in ECN90 β cells. The first three columns indicate the source protein (UniProt accession number and gene name) and whether its expression is β-cell-enriched. Subsequent columns indicate predicted HLA-I restrictions and ranks based on NetMHCpan4.1a; the condition in which the peptides were identified; the full peptide sequence; its N-terminal and C-terminal splicing fragments; the forward or reverse splicing direction; and the inter-fragment distance. The last two columns indicate whether the peptide was ultimately validated by spectral matching; and the Ensembl accession number when mRNA splice isoforms were also assigned by MARS as an alternative aa sequence match.

UniProt	Gene name	β-cell enriched expression	Predicted HLA restriction	Rank	Presence	Sequence	N-terminal fragment	C-terminal fragment	Splicing direction	Inter-fragment distance	Spectral matching	Ensembl match
P46821	MAP1B	Yes	A03:01	0.03	Both	ATAGPGTVK	VK	ATAGPGT	Reverse	25	No	
Q9BTF0	THUMPD2	Yes	B49:01/B40:01	0.20/0.50	IFNa	EEVGKVIGI	EVGKVIGI	E	Reverse	17	No	
							E	EVGKVIGI	Forward	38	No	
							EE	VGKVIGI	Forward	38	No	
O75592	MYCBP2	Yes	A03:01	0.14	IFNa	GSINLSMLK	GSIN	LSMLK	Forward	4	Yes	
Q9Y5W8	SNX13	No	A03:01	0.14	IFNa	GSINLSMLK	MLK	GSINLS	Reverse	20	Yes	
P52292	KPNA2	Yes	A03:01	0.02	IFNa	GTNPKTNIQK	TNPKTNIQK	G	Reverse	23	Yes	
							G	TNPKTNIQK	Forward	26	Yes	
							GT	NPKTNIQK	Forward	21	Yes	
Q9ULU8	CADPS	Yes	B49:01/B40:01	0.06-0.08/0.45-0.50	IFNa	IEENVGRLI	IE	IEENVGRLI	Forward	16	Yes	ENSP00000484365.1
							IEE	NVGRLI	Forward	6	Yes	ENSP00000484365.1
							IEEN	VGRLI	Forward	6	Yes	ENSP00000484365.1
					IFNa	LEENVGRLI	LE	ENVGRLI	Forward	23	Yes	ENSP00000484365.1
Q8TDR0	TRAF3IP1	No	B49:01/B40:01	0.34/1.30	Both	KEIEKGKI	KEIEK	GKI	Forward	15	Yes	
Q8IXS2	CCDC65	No	B49:01/B40:01	0.34/1.30	Both		KEIEK	GKI	Forward	16	Yes	
Q8WXH0	SYNE2	No	B40:01/B49:01	0.45/0.84	Both	KEIEKGKL	IEKGKL	KE	Reverse	27	Yes	
P46821	MAP1B	Yes	B40:01/B49:01	0.69/1.24	Both	KELEKGKI	EKGKI	KEL	Reverse	38	Yes	
							KGKI	KELE	Reverse	4	Yes	
Q96AE7	TTC17	No	B49:01/B40:01	0.34/1.30	Both	KELEKGKI	GKI	KELEK	Reverse	29	Yes	
P0DMV9	HS71B	No	B49:01/B40:01	0.34/1.30	Both	KELEKGKI	KGKI	KELE	Reverse	34	Yes	
P11142	HSP7	No			Both	KELEKGKI	GKI	KELEK	Reverse	34	Yes	
P0DMV8	HS71A	No			Both	KELEKGKI	KGKI	KELE	Reverse	34	Yes	
P34931	HS71L	No			Both	KELEKGKI	KGKI	KELE	Reverse	34	Yes	
Q16181	SETPIN7	No			Both	KELEKGKI	KELE	KGKI	Forward	34	Yes	
Q16851	UGPA	No			Both	KELEKGKI	KELEK	GKI	Forward	37	Yes	
Q12802	AKP13	No	B40:01/B49:01	0.74/1.28	Both	KELEKGKL	EKGKL	KEL	Reverse	30	Yes	
Q8WY36	BBX	No			Both	KELEKGKL	KL	KELEKG	Reverse	11	Yes	
							L	KELEKGK	Reverse	31	Yes	
							KELEKGK	L	Forward	40	Yes	
Q9NZR1	TMOD2	No			Both	KELEKGKL	KELEK	GKL	Forward	10	Yes	
Q9UPW8	UNC13A	Yes	A03:01	0.01	Basal	KLITLPSIK	PSIK	KLITL	Reverse	6	Yes	
A4D1F6	LRRD1	Yes	A03:01	0.01	Basal	KLITLPSLK	TLPSLK	KLI	Reverse	23	Yes	
Q8IWU4	SLC30A8	Yes	A03:01	0.01	Both	KMYAFTLER	LER	KMYAFT	Reverse	9	Yes	
							ER	KMYAFTL	Reverse	9	Yes	
							R	KMYAFTLE	Reverse	9	Yes	
							KMYAFTL	ER	Forward	17	Yes	
							KMYAFTLE	R	Forward	17	Yes	
P11021	HSPA5	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	33	Yes	
							RIINEPTAA	K	Forward	12	Yes	
P17066	HSP76	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	30	Yes	
							RIINEPTAA	K	Forward	40	Yes	
P0DMV9	HS71B	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	11	Yes	
							RIINEPTAA	K	Forward	40	Yes	
P54652	HSP72	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	32	Yes	
							RIINEPTAA	K	Forward	15	Yes	
P11142	HSP7C	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	33	Yes	
							RIINEPTAA	K	Forward	40	Yes	
P0DMV8	HS71A	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	11	Yes	
							RIINEPTAA	K	Forward	40	Yes	
P34931	HS71L	No	A03:01	0.01	IFNa	RIINEPTAAK	K	RIINEPTAA	Reverse	11	Yes	
							RIINEPTAA	K	Forward	40	Yes	
Q9BTF0	THUMPD2	Yes	B49:01/B40:01	0.02/0.14	IFNa	SEVGKVIGI	EVGKVIGI	S	Reverse	40	Yes	
							S	EVGKVIGI	Forward	14	Yes	
Q12769	NU160	No	B40:01/B49:01	0.02/0.11	IFNa	SEVGKVIGL	SEVGK	VIGL	Forward	32	Yes	
P50747	BPL1	No	B40:01/B49:01	0.03/0.11	IFNa	SEVGKVLGL	GKVLGL	SEV	Reverse	31	Yes	
Q8TDD5	MCLN3	No	A02:01	0.07	IFNa	TIIGSILKV	V	TIIGSILK	Reverse	33	Yes	
							TIIGSILK	V	Forward	28	Yes	
Q8IZK6	MCLN2	No	A02:01	0.07	IFNa	TIIGSILKV	V	TIIGSILK	Reverse	22	Yes	
							TIIGSILK	V	Forward	40	Yes	
Q8NGH8	OR56A4	No	A02:01	0.07	IFNa	TLLGSILKV	TLLGS	ILKV	Forward	11	Yes	
Q2V2M9	FHOD3	No	A02:01	0.07	IFNa	TLIGSLIKV	LIKV	TLIGS	Reverse	40	Yes	
O15085	ARHGEF11	Yes	A02:01	0.02	IFNa	TLLGSIIKV	IIKV	TLLGS	Reverse	25	Yes	

Table S5. Predicted HLA-A2 and -A3-restricted peptides identified in ECN90 β cells and primary human islets. The comparison was performed between the conventional peptides eluted from ECN90 β cells and from two islet samples. For PTMs defined as likely artifactual, the native peptide sequence was retained. From left to right, columns indicated: gene name of the source protein, peptide sequence, the treatment condition of ECN90 β cells in which the peptide was identified, aa length, HLA-I restriction and rank (based on the HLA-I haplotypes of ECN90 β cells) and the presence in the immunopeptidome of islet sample 1 and 2.

Gene name	Peptide	Presence	Length	Granule protein	HLA restriction	Rank	Islet 1	Islet 2
CHGA	GLSAEPGWQA	Basal	10	Yes	A02:01	0.8389	No	Yes
MAP1B	SLYASLTSEK	IFN-α	10	No	A03:01	0.0189	Yes	No
KIF1A	AVFNIIFTQK	Both	10	No	A03:01	0.022	Yes	No
KIF1A	LLYPVPLVH	Both	9	No	A03:01	0.0638	Yes	Yes
KIF1A	VLDTSVAYV	IFN-α	9	No	A02:01	0.0238	Yes	Yes
KIF1A	KLSEMSVTL	Both	9	No	A02:01	0.0117	Yes	Yes
RPS17	KLLDFGSLSNL	Both	11	No	A02:01	0.2199	Yes	No
RFX6	KTITQIVKDK	Both	10	No	A03:01	0.0871	Yes	Yes
RFX6	SVYSGKGLTR	Both	10	No	A03:01	0.086	Yes	Yes
HSPA5	KLSLVAAML	Both	9	No	A02:01	0.8236	Yes	Yes
HSPA5	KVYEGERPLTK	IFN-α	11	No	A03:01	0.0031	Yes	Yes
ABCC8	IIFESPFNK	Both	9	Yes	A03:01	0.01	No	Yes
ABCC8	GTIQREGTLK	Both	10	Yes	A03:01	0.3892	No	Yes
DHX15	ILMGVLKEV	Both	9	No	A02:01	0.0154	Yes	Yes
CPE	ASAPGYLAITK	Both	11	Yes	A03:01	0.2238	No	Yes
TTR	STTAVVTNPK	Both	10	No	A03:01	0.577	Yes	Yes
PTPRN	SLADVTQQA	IFN-α	9	Yes	A02:01	0.0469	Yes	Yes
PTPRN	SLYHVYEV	Both	8	Yes	A02:01	0.2596	Yes	No
ORC3	KTDPKMLSK	Both	9	No	A03:01	0.0074	Yes	No
SLC30A8	RTYLVNDKAAK	Both	11	Yes	A03:01	0.1627	Yes	Yes
UCHL1	GTIGLIHAV	Both	9	No	A02:01	0.1773	Yes	Yes
COPS2	SLNQAVVSK	Both	9	No	A03:01	0.0091	Yes	Yes
RTN1	STYFTGILQK	Both	10	No	A03:01	0.0091	Yes	No
PAX6	GLISPGVSV	Basal	9	No	A02:01	0.0876	Yes	No
PPP4R1	TLLEKDYFL	Both	9	No	A02:01	0.0229	Yes	Yes
UCN3	MLMPVHFLL	IFN-α	9	Yes	A02:01	0.0379	Yes	Yes
CTNNA1	ASYVASTKYQK	Basal	11	No	A03:01	0.2798	Yes	Yes
CTNNA1	KMKAPEKKPLVK	Basal	12	No	A03:01	0.272	Yes	Yes
DMXL1	GLFEAAHVL	IFN-α	9	No	A02:01	0.0336	No	Yes
EPB41L3	SMYGVDLHHA	Both	10	No	A02:01	0.3101	No	Yes
EPB41L3	NLLEKDYFGL	IFN-α	10	No	A02:01	0.7314	Yes	No
HNRNPCL4	SVHKGFAFV	Both	9	No	A02:01	0.8047	Yes	Yes
HNRNPCL4	SLLENLEKI	Both	9	No	A02:01	0.0209	Yes	Yes
INS	ALWGPDPAAA	Both	10	Yes	A02:01	0.3268	Yes	Yes
INS	HLVEALYLV	Both	9	Yes	A02:01	0.0364	Yes	Yes
INS	ALWGPDPAA	Basal	9	Yes	A02:01	0.2955	Yes	Yes
NDST4	GLFTHTIFY	IFN-α	9	No	A03:01	0.0895	Yes	Yes
PTPRN	AVAEEVNAILK	Both	11	Yes	A03:01	0.4261	Yes	No
RGPD5	KLFPGSPAIY	Both	10	No	A03:01	0.0969	Yes	Yes
RXRA	KLIGDTPIDTFL	IFN-α	12	No	A02:01	0.4879	Yes	No

Table S6. Putative HLA-E ligands in the immunopeptidome of ECN90 β cells. To assign HLA-E restriction, a ≥3-fold NetMHC4.1a rank score difference with the second-best allele was used. This is a less stringent criterion than the one used for assigning HLA-E restriction in Tables S1-4, namely a predicted HLA-E restriction without any other HLA-I allele with a ranking score <2. Both conventional peptides and peptides carrying PTMs (listed as native sequences, as all PTMs were defined as likely artifactual) were used. Legend is the same as in Table S2.

Peptide	Presence	Basal - # of replicates	IFN-α - # of replicates	Length	UniProt	Position	Gene name	Gene synonyms	Gene description	Granule protein	HLA restriction	Rank
VTKEGLEL	IFN-α	0/4	2/4	8	P08238;P079	536-543	HSP90AB1;HSP	90AA1;HSP90AB3	Heat shock protein	N	C03:04/E01:01	0.6051/1.490
RSFEGREL	Both	2/4	2/4	8	P16870	81-88	CPE		Carboxypeptidase E	Y	C03:04/E01:01	0.4425/1.189
KLGPALKL	IFN-α	0/4	2/4	8	Q9UQR0;Q96	680-687	SCML2;SCMH	Scml3	Sex comb on midleg	N	E01:01	1.5466
KIYNKIMHL	IFN-α	0/4	1/4	9	Q09428	394-402	ABCC8	ABC36,HHF1,HI,HR	ATP binding cassett	Y	A02:01/E01:01	0.0403/0.0147
ATDPTIHTV	IFN-α	0/4	1/4	9	Q9H6U8	450-458	ALG9	DIBD1	ALG9, alpha-1,2-ma	N	A02:01/E01:01	0.0376/0.0290
KMPLKPIFL	Both	1/4	2/4	9	Q5TH69	1407-1415	ARFGEF3	BIG3,C6orf92,KIAA	ARFGEF family mem	N	E01:01	0.0042
ASNSVILHL	Both	2/4	4/4	9	O75973;Q862	216-224	C1QL1;C1QL4	C1QRF,C1QTNF11,	Complement C1q lik	N	C07:01/E01:01	0.2929/0.1136
ISEEGFHYL	IFN-α	0/4	3/4	9	Q13555;Q135	80-88	CAMK2G;CAM	CAMKG,CAMKD,CA	Calcium/calmodulin	N	C03:04/E01:01	0.0974/0.0769
RSAAVLALL	Both	2/4	3/4	9	P10645	44471	CHGA		Chromogranin A	Y	E01:01	0.2083
KMDQLAKEL	Both	4/4	4/4	9	P10645	344-352	CHGA		Chromogranin A	Y	E01:01	0.0406
GFRGPGPQL	Both	4/4	4/4	9	P10645	382-390	CHGA		Chromogranin A	Y	C07:01/E01:01	1.3265/1.3126
KVAHQLQAL	Both	1/4	1/4	9	P10645	446-454	CHGA		Chromogranin A	Y	C03:04/E01:01	0.487/0.5707
RSFEGRELL	IFN-α	0/4	1/4	9	P16870	81-89	CPE		Carboxypeptidase E	Y	E01:01	0.0736
KMMSETLNF	Both	4/4	4/4	9	P16870	468-476	CPE		Carboxypeptidase E	Y	E01:01	0.1764
AGHAGHTSL	Basal	1/4	0/4	9	O43143	103-111	DHX15	DBP1,DDX15,HRH2	DEAH-box helicase	N	C03:04/E01:01	0.5691/0.4973
KTAPPGSAV	IFN-α	0/4	2/4	9	Q9BTC0	549-557	DIDO1	BYE1,C20orf158,D	Death inducer-oblit	N	C03:04/E01:01	0.268/0.3481
VMPEAVPEV	IFN-α	0/4	1/4	9	Q9BTC0	902-910	DIDO1	BYE1,C20orf158,D	Death inducer-oblit	N	A02:01/E01:01	0.0088/0.0134
YTSKLVKTL	IFN-α	0/4	2/4	9	Q7RTS9	210-218	DYM	DMC,FLJ20071,SM	Dymeclin	N	C03:04/E01:01	0.0832/0.1081
VSLDRYLAI	IFN-α	0/4	1/4	9	O43603	119-127	GALR2	GALNR2	Galanin receptor 2	N	C03:04/E01:01	0.6528/0.3132
SVSQINHKL	IFN-α	0/4	2/4	9	Q9HD26	86-94	GOPC	CAL,dJ94G16.2,FIG	Golgi associated PD	N	C03:04/E01:01	0.2668/0.1199
RVPPPPPIA	Both	3/4	2/4	9	P0DMR1;B7Z	129-137	HNRNPCL4;HI	HNRNPCP5,HNRPC	Heterogeneous nuc	N	A02:01/E01:01	1.3086/0.4901
KLSLVAAML	Both	3/4	3/4	9	P11021	44471	HSPA5	BiP,GRP78	Heat shock protein	N	A02:01/E01:01	0.8236/0.7302
RVMEHFIKL	IFN-α	0/4	1/4	9	P11021	261-269	HSPA5	BiP,GRP78	Heat shock protein	N	A02:01/E01:01	0.0432/0.0157
KLSEMSVTL	Both	3/4	4/4	9	Q12756	1480-1488	KIF1A	ATSV,C2orf20,SPG	Kinesin family mem	N	A02:01/E01:01	0.0117/0.0197
GTVEQIVYL	IFN-α	0/4	1/4	9	P52292	408-416	KPNA2	IPOA1,QIP2,RCH1,	Karyopherin subunit	N	A02:01/E01:01	0.2869/0.6650
KTVDGPREL	Both	2/4	3/4	9	P16519	276-284	PCSK2	NEC2,PC2,SPC2	Proprotein converta	Y	C03:04/E01:01	0.0747/0.1978
RLYQEVHRL	IFN-α	0/4	3/4	9	Q92932	387-395	PTPRN2	IA-2beta,ICAAR,KI/	Protein tyrosine pho	N	A02:01/E01:01	0.0126/0.0274
RLSDTPPLL	IFN-α	0/4	3/4	9	P20337	205-213	RAB3B		RAB3B, member RA	N	A02:01/E01:01	0.023/0.0142
GLKDFKTFL	IFN-α	0/4	2/4	9	P49792;Q996	2176-2184	RANBP2;RGPD	ADANE,ANE1,NUP	RAN binding protein	N	A02:01/E01:01	0.381/1.0707
RASGQAFEL	Both	2/4	4/4	9	Q93045;P169	48-56	STMN2;STMN	C1orf215,FLJ3220(	Stathmin 1;Stathmir	N	C03:04/E01:01	0.1447/0.3949
RLAEQIATL	Basal	1/4	0/4	9	P61764	171-179	STXBP1	hUNC18,MUNC18	Syntaxin binding pro	N	A02:01/E01:01	0.0108/0.0228
VSQEVTRSL	IFN-α	0/4	3/4	9	P61764	299-307	STXBP1	hUNC18,MUNC18	Syntaxin binding pro	N	C03:04/E01:01	0.0348/0.0948
RISEQTYQL	IFN-α	0/4	3/4	9	P61764	467-475	STXBP1	hUNC18,MUNC18	Syntaxin binding pro	N	A02:01/E01:01	0.1205/0.0799
YQNDKITQL	IFN-α	0/4	1/4	9	Q13207;O151	266-274	TBX2;TBX3	TBX3-ISO,UMS,XH	T-box 2;T-box 3	N	C07:01/E01:01	0.0175/0.0136
MASHRLLLL	IFN-α	0/4	1/4	9	P02766	44440	TTR	CTS,CTS1,HsT2651	Transthyretin	N	C03:04/E01:01	0.1675/0.3021
SYIREQYVL	IFN-α	0/4	4/4	9	Q8N398	659-667	VWA5B2	DKFZp761K032,LO	Von Willebrand fact	N	C07:01/E01:01	0.5848/0.7440
YGFRGPGPQL	Both	4/4	4/4	10	P10645	381-390	CHGA		Chromogranin A	Y	C03:04/E01:01	0.4651/0.934
ATSPIIIHRL	IFN-α	0/4	1/4	10	Q9UBD5	253-262	ORC3	IMAGE50150,LATH	Origin recognition c	N	E01:01	0.3748
KQMEQISQFL	IFN-α	0/4	4/4	10	P37802;Q9UI	88-97	TAGLN2;TAGL	HA1756,KIAA0120,	Transgelin 2;Transg	N	A02:01/E01:01	0.3529/0.279
RAYGFRGPGPQL	Both	4/4	4/4	12	P10645	379-390	CHGA		Chromogranin A	Y	C03:04/E01:01	0.8606/1.570
GIPPAPRGV	Basal	3/4	0/4	9	P11021;P111	486-494	HSPA5;HSPA8	HSC70,HSC71,HSP	Heat shock protein	N	A02:01/E01:01	1.5107/1.423

Table S7. PTMs identified in the immunopeptidome of ECN90 β cells. The first two columns indicate the PTM name and mass shift observed in MS. The next three columns list the number of peptides and source proteins identified with the indicated PTM, along with their distribution between the two conditions. Specific PTM searches were run for those PTMs meeting our two selection criteria, i.e. enrichment in either basal or IFN-α-treated condition and PTM introducible on synthetic peptides (light grey). PTMs detected in very few (<10) peptides were only retained if found in predicted HLA-A2or -A3 binders. In total, 253 PTMs on 247 peptides (with 6 peptides carrying two distinct PTMs) were identified. The last column indicates the number of peptides for which the PTM was validated by spectral matching, i.e. a match with the peptide synthesized with the modification, but not in the peptide synthesized in its native form. Given their unlikely artifactual origin, phosphorylated and glutathionylated peptides were not compared to the unmodified sequence. Modified peptides validated as likely biological are indicated in bold.

PTM name	Mass shift (Da)	# of peptides	# of proteins	Basal - IFNa (%)	Spectral Matching
2-amino-3-oxo-butanoic_acid	-2.02	1	1	0 - 100	ND
Acetylation (K)	+42.01	2	2	0 - 100	0/2
Acetylation (N-term)	+42.01	6	3	38 - 62	ND
Acetylation (Protein N-term)	+42.01	3	2	66 - 34	ND
Acetylation (TSCYH)	+42.01	1	1	50 - 50	ND
Addition of Carbon to cysteine	+12.00	1	1	100 - 0	ND
Amidation	-0.98	2	2	50 - 50	ND
Aminoethylbenzenesulfonylation	+183.04	1	1	50 - 50	ND
Carbamidomethylation (DHKE X@N-term)	+57.02	4	4	50 - 50	ND
Carboxylation (DKW)	+43.99	1	1	50 - 50	ND
Crotonaldehyde	+70.04	7	2	50 - 50	ND
Cysteine oxidation to cysteic acid	+47.98	7	1	38 - 62	ND
Cysteinylation	+119.00	8	3	60 - 40	**1/1**
Deamidation (NQ)	+0.98	55	14	48 - 32	0/13
Dehydration	-18.01	6	4	29 - 71	ND
Dihydroxy	+31.99	3	2	40 - 60	ND
Dimethylation(KR)	+28.03	1	1	50 - 50	ND
Formylation (Protein N-term)	+27.99	1	1	0 - 100	ND
Glutathione disulfide	+305.07	37	11	55 - 45	**8/9**
Methylation(C-term)	+14.02	1	1	0 - 100	ND
Methylation(KR)	+14.02	2	2	50 - 50	ND
Oxidation (HW)	+15.99	2	2	33 - 66	ND
Oxidation (M)	+15.99	64	27	50 - 50	ND
Phospho-propargylamine	+117.00	1	1	100 - 0	ND
Phosphorylation (STY)	+19.97	8	4	75 - 25	**1/1**
Pyro-glu from Q	-17.03	5	4	67 - 33	ND
Replacement of 2 protons by calcium	+37.95	1	1	50 - 50	ND
Replacement of 2 protons by magnesium	+21.97	1	1	50 - 50	ND
Replacement of proton with ammonium ion	+17.03	1	1	50 - 50	ND
Sodium adduct	+21.98	1	1	50 - 50	ND
Sulfation	+79.96	3	1	100 - 0	ND
Sulphone	+31.99	14	8	45 - 55	**2/5**
Tryptophan oxidation to kynurenin	+3.99	1	1	50 - 50	0/1
Tryptophan oxidation to oxolactone	+13.98	1	1	50 - 50	ND

Table S8. Summary of CD8$^+$ T-cell islet epitopes identified. The first columns detail the identity of each peptide. MMr$^+$ donors were defined as those with MMr$^+$CD8$^+$ T-cell frequencies >10^{-6} and ≥3 MMr$^+$CD8$^+$ T cells counted. Median MMr$^+$CD8$^+$ T-cell frequencies, MMr$^+$ counts and percent effector/memory fractions are listed in subsequent columns. Percent effector/memory fractions were not assigned (NA) when <5 MMr$^+$ cells were counted. Validated epitopes are displayed in red fonts and shown in Fig. 5. Bold horizontal lines identify the 5 distinct epitope panels used for MMr assays.

Peptide group	Protein	Position	Sequence	MMr+ donors	Frequency median	MMr+ counts median (range)	% effector/memory median (range)
Granule	ABCC8	394-402	KIYNKIMHL	4/5	8,55E-06	5 (2-185)	33 (20-43)
Granule	ABCC8	862-870	HLSDHLMQA	0/5	0	0 (0-0)	NA
Granule	ABCC8	1259-1267	VLIAAVTSI	1/5	9,50E-07	1 (0-4)	NA
β-enriched	BEX1	100-108	QLSHSLRAV	4/5	5,60E-06	5 (2-13)	35 (20-62)
β-enriched	CADPS	680-688	SLFEMVQRL	5/5	2,00E-05	18 (7-21)	86 (29-95)
β-enriched	CADPS	730-738	YLRDLLERA	0/5	0	0 (0-0)	NA
Granule	CPE	71-79	TAISRIYTV	1/5	1,60E-06	1 (0-5)	0
Granule	CPE	412-421	RLLIPGNYKL	1/5	9,50E-07	1 (0-5)	40
β-enriched	CRMP1	374-382	KMDENQFVA	1/5	1,90E-06	2 (1-4)	NA
β-enriched	DACH1	189-198	KMVDLRGAKV	5/5	3,40E-05	24 (8-66)	11 (7-21)
β-enriched	DACH1	221-229	HLVGGLHTV	0/5	0	0 (0-0)	NA
β-enriched	DPYSL2	497-506	GLYDGPVCEV	0/5	0	0 (0-0)	NA
β-enriched	ELAVL3	156-164	ILVDQVTGV	5/5	6,70E-05	43 (13-419)	12 (4-39)
β-enriched	EPB41L2	144-153	NLLEKDYFGL	2/5	2,80E-06	2 (0-5)	0
β-enriched	EPB41L2	402-411	SMYGVDLHHA	0/5	0	0 (0-0)	NA
β-enriched	FRYL	1003-1011	GLDNETHFL	0/5	8,20E-07	1 (0-2)	NA
β-enriched	HNRNPCL4	129-137	RVPPPPPIA	0/5	0	0 (0-0)	NA
β-enriched	HNRNPCL4	196-204	SLLENLEKI	3/5	3,70E-06	3 (0-9)	31 (22-40)
β-enriched	HSPA5	2-10	KLSLVAAML	4/5	3,30E-06	4 (1-6)	53 (40-67)
β-enriched	HSPA5	197-205	RIINEPTAA	4/5	1,30E-05	13 (1-28)	77 (69-88)
β-enriched	HSPA5	486-494	GIPPAPRGV	0/5	0	0 (0-0)	NA
β-enriched	JAKMIP2	29-37	KLTDIQIEL	0/5	0	0 (0-2)	NA
β-enriched	KIAA1324	761-769	SLADRLIGV	1/5	1,50E-06	2 (1-4)	NA
β-enriched	KIAA1324	903-911	KTIDFWLKV	3/5	3,60E-06	5 (0-6)	0 (0-60)
Granule-assoc	KIF1A	419-427	SLHERILFA	5/5	2,20E-05	21 (8-51)	25 (6-38)
Granule-assoc	KIF1A	510-518	YIKDGITRV	0/5	0	0 (0-1)	NA
Granule-assoc	KIF1A	1347-1355	VLDTSVAYV	0/5	0	0 (0-0)	NA
β-enriched	NEURL1	435-443	GLHGTITQI	5/5	9,00E-06	11 (3-43)	20 (17-47)
β-enriched	PAX6	395-403	GLISPGVSV	1/5	8,20E-07	1 (0-3)	NA
Granule	PCSK1N	24-33	GLFRPPPALC	0/5	ND	0 (0-1)	NA
β-enriched	PEG10	36-44	KLTEENTTL	2/5	1,70E-06	2 (0-5)	60
β-enriched	PEG10	229-237	ALIGQCIHI	0/5	0	0 (0-0)	NA
β-enriched	PLCB4	383-391	ILFKDVIQA	1/5	1,85E-06	1 (0-4)	NA
Granule	PTPRN	57-66	GLFGQCQVGV	3/5	2,60E-06	3 (0-4)	NA
Granule	PTPRN	526-534	SLADVTQQA	4/5	8,40E-06	5 (1-13)	52 (20-60)
β-enriched	PTPRN2	387-395	RLYQEVHRL	0/5	0	0 (0-0)	NA
β-enriched	RAB3B	205-213	RLSDTPPLL	3/5	2,80E-06	3 (1-3)	NA
β-enriched	RAP1GAP2	459-468	ALLDNLHDEL	1/5	2,10E-06	2 (0-3)	NA
β-enriched	RFX6	408-416	SMVSDIERV	1/5	1,90E-06	2 (0-8)	75
β-enriched	RGPD5	1681-1689	VLMEQIKLL	5/5	2,40E-05	17 (11-72)	27 (6-36)
β-enriched	RXRA	440-451	KLIGDTPIDTFL	0/5	0	0 (0-0)	NA
Granule	SCG5	149-160	HLFDPEHDYPGL	0/5	0	0 (0-1)	NA
Granule	SCG5	186-195	YLQGQRLDNV	5/5	1,60E-04	126 (61-304)	11 (4-25)
β-enriched	SLC7A8	304-313	VTFGEKLLGV	2/5	4,20E-06	2 (0-6)	18 (17-20)
β-enriched	SLC7A8	441-449	SLWSEPVVC	0/5	0	0 (0-1)	NA
β-enriched	SNCA	47-55	GVVHGVATV	2/5	1,50E-06	1 (0-51)	78
β-enriched	STXBP1	171-179	RLAEQIATL	0/5	0	0 (0-0)	NA
β-enriched	STXBP1	295-303	HIAEVSQEV	0/5	9,60E-07	1 (0-2)	NA
β-enriched	STXBP1	467-475	RISEQTYQL	2/5	3,00E-06	2 (0-4)	NA
β-enriched	TAGLN2	88-97	KQMEQISQFL	5/5	1,30E-05	11 (4-21)	58 (20-93)
β-enriched	THUMPD2	76-85	KQFPLIISSV	0/5	0	0 (0-0)	NA
β-enriched	TMEM132D	737-745	SLDEKVVSI	0/5	0	0 (0-1)	NA
β-enriched	TMOD1	263-272	FISGAGILRL	2/5	1,80E-06	2 (0-5)	80
Granule	UCN3	1-9	MLMPVHFLL	0/5	0	0 (0-2)	NA
β-enriched	UNC13A	825-833	NLFHFVTDV	5/5	3,30E-05	23 (7-37)	19 (0-41)
β-enriched	VWA5B2	105-113	HLVLDLAQA	0/5	0	0 (0-0)	NA
Reference	R-GAD2	114-122	VMNILLQYV	0/5	0	0 (0-0)	NA
Reference	R-GAD2	141-149	LLQEYNWEL	5/5	4,10E-05	39 (14-50)	21 (8-33)
Reference	R-INS	6-14	RLLPLLALL	4/5	6,70E-06	4 (2-9)	22 (11-33)
Reference	R-INS	90-99	GIVEQCCTSI	0/5	0	0 (0-0)	NA
Reference	R-INS-DRiP	1-9	MLYQHLLPL	0/5	8,60E-07	1 (0-2)	NA
Granule	R-PTPRN	790-798	TIADFWQMV	0/5	0	0 (0-0)	NA
Reference	R-PTPRN	805-813	VIVMLTPLV	0/5	0	0 (0-1)	NA
Reference	R-PTPRN	830-839	SLYHVYEVNL	0/5	1,00E-06	1 (0-2)	NA
Reference	R-PTPRN	962-970	ALTAVAEEV	5/5	2,20E-04	198 (104-435)	11 (8-27)
Reference	R-SCG5-009	186-194	FLSGAVNRL	0/5	0	0 (0-0)	NA
Reference	R-SLC30A8	186-194	VAANIVLTV	0/5	0	0 (0-2)	NA
*β-enriched	CTNNA1	111-119	FADDPC(glut)SSV	0/5	1,44E-06	1 (0-2)	NA
*β-enriched	SLC7A8	441-449	SLWSEPVVC(glut)	0/5	0	0 (0-0)	NA
*Granule	PCSK1N	24-33	GLFRPPPALC(glut)	5/5	5,60E-05	44 (18-513)	13 (2-33)
*β-enriched	PEG10	229-237	ALIGQC(glut)IHI	0/5	0	0 (0-0)	NA
*Granule	PTPRN	57-66	GLFGQC(glut)QVGV	1/5	0	0 (0-3)	NA
Control	CMV pp65	495-503	NLVPMVATV	5/5	1,60E-04	154 (24-5466)	13 (7-98)
Control	Flu MP	58-66	GILGFVFTL	5/5	3,90E-04	398 (187-738)	98 (96-100)
Control	MelanA, ELA	26-35	ELAGIGILTV	0/5	0	0 (0-2)	NA

Table S9. Primers designed for proteasome/immunoproteasome gene expression

Gene	Forward	Reverse
PMSB5	GCCTTCAAGTTCCGCCAT	TGCCTAGCAGGTATGGGTTG
PMSB6	CAACCACTGGGTCCTACATCG	GGAAACCGAGCTGGTAGGTG
PMSB7	AAGTTGCCTTATGTCACCATGG	ATGGCTTCGCTCACCAGAT
PMSB8	GGCTGTACTATCTGCGAAATGG	AGTCCAGGACCCTTCTTATCCC
PMSB9	ATGCTGACTCGACAGCCTTT	GAGCAATAGCGTCTGTGGTG
PMSB10	ACTGCCAAAGAAATGCATCAT	ATCGTTAGTGGCTCGCGTAT
ACTB	ACTCTTCCAGCCTTCCTTCC	CGTACAGGTCTTTGCGGATG
GAPDH	TCGTGGAAGGACTCATGACC	ATGATGTTCTGGAGAGCCCC

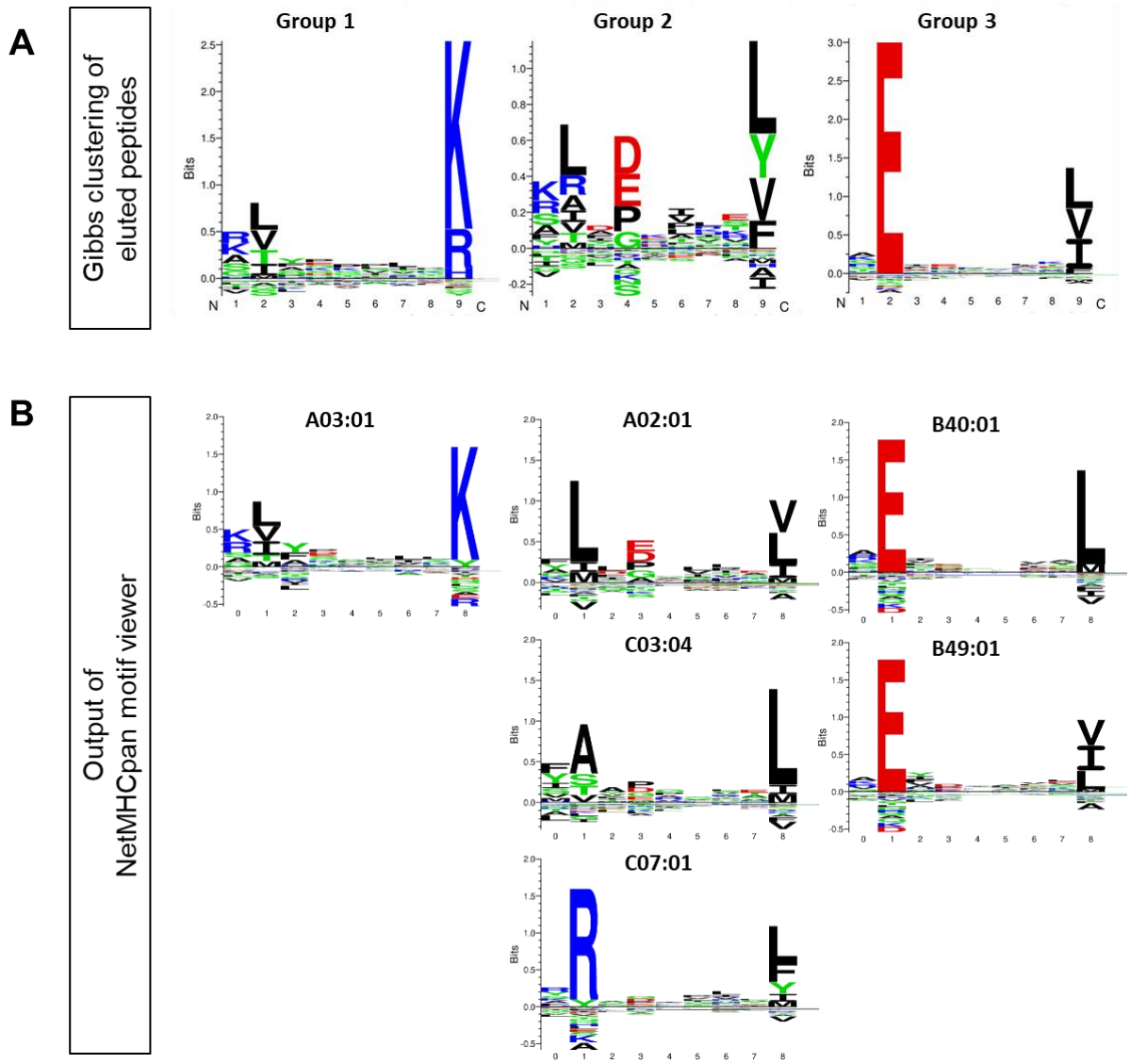


Figure S1. Gibbs clustering of 8-14mer peptides. A. All unique conventional and post-translationally modified candidates identified in the 8-14 aa length range were clustered using GibbsCluster. The *x* axis indicates the residue position within the 9mer core sequence. Each aa is represented by its single letter code, with its size proportional to its frequency at the indicated position. All peptides were inputed as unmodified sequences. **B.** NetMHCpan4.1a sequence motifs for the alleles expressed by the ECN90 β cells.

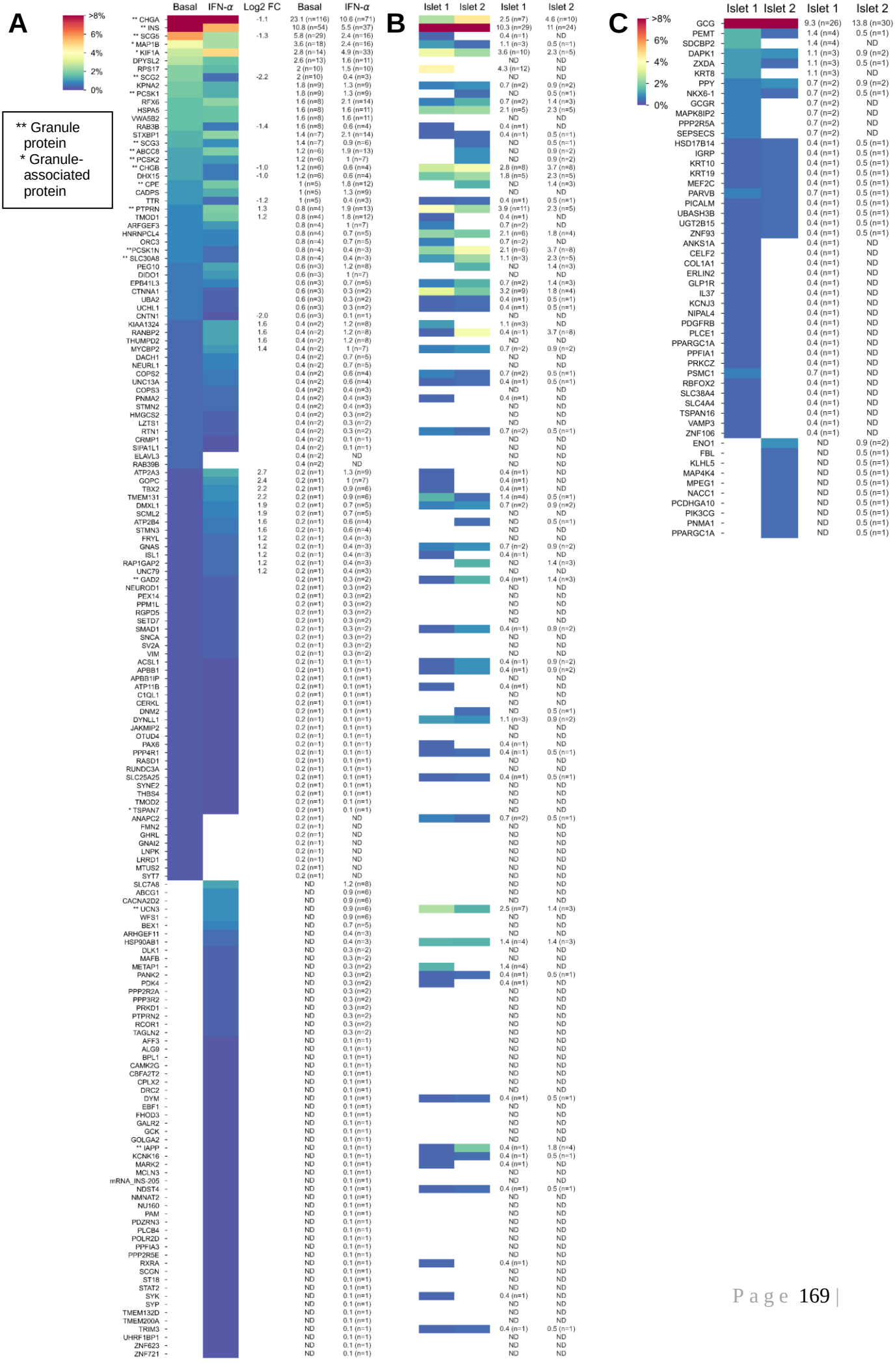

A
>8%
6%
4%
2%
0%
** Granule protein
* Granule-associated protein
Basal
IFN-α
Log2 FC
Basal
IFN-α
** CHGA
** INS
** SCG5
* MAP1B
* KIF1A
DPYSL2
RPS17
** SCG2
KPNA2
** PCSK1
RFX6
HSPA5
VWA5B2
RAB3B
STXBP1
** SCG3
** ABCC8
** PCSK2
** CHGB
DHX15
** CPE
CADPS
TTR
** PTPRN
TMOD1
ARFGEF3
HNRNPCL4
ORC3
**PCSK1N
** SLC30A8
PEG10
DIDO1
EPB41L3
CTNNA1
UBA2
UCHL1
CNTN1
KIAA1324
RANBP2
THUMPD2
MYCBP2
DACH1
NEURL1
COPS2
UNC13A
COPS3
PNMA2
STMN2
HMGCS2
LZTS1
RTN1
CRMP1
SIPA1L1
ELAVL3
RAB39B
ATP2A3
GOPC
TBX2
TMEM131
DMXL1
SCML2
ATP2B4
STMN3
FRYL
GNAS
ISL1
RAP1GAP2
UNC79
** GAD2
NEUROD1
PEX14
PPM1L
RGPD5
SETD7
SMAD1
SNCA
SV2A
VIM
ACSL1
APBB1
APBB1IP
ATP1B
C1QL1
CERKL
DNM2
DYNLL1
JAKMIP2
OTUD4
PAX6
PPP4R1
RASD1
RUNDC3A
SLC25A25
SYNE2
THBS4
TMOD2
* TSPAN7
ANAPC2
FMN2
GHRL
GNAI2
LNPK
LRRD1
MTUS2
SYT7
SLC7A8
ABCG1
CACNA2D2
** UCN3
WFS1
BEX1
ARHGEF11
HSP90AB1
DLK1
MAFB
METAP1
PANK2
PDK4
PPP2R2A
PPP3R2
PRKD1
PTPRN2
RCOR1
TAGLN2
AFF3
ALG9
BPL1
CAMK2G
CBFA2T2
CPLX2
DRC2
DYM
EBF1
FHOD3
GALR2
GCK
GOLGA2
** IAPP
KCNK16
MARK2
MCLN3
mRNA_INS-205
NDST4
NMNAT2
NU160
PAM
PDZRN3
PLCB4
POLR2D
PPFIA3
PPP2R5E
RXRA
SCGN
ST18
STAT2
SYK
SYP
TMEM132D
TMEM200A
TRIM3
UHRF1BP1
ZNF623
ZNF721
-1.1
-1.3
-2.2
-1.4
-1.0
-1.0
-1.2
1.3
1.2
-2.0
1.6
1.6
1.6
1.4
2.7
2.4
2.2
2.2
1.9
1.9
1.6
1.6
1.2
1.2
1.2
1.2
1.2
23.1 (n=116)
10.8 (n=54)
5.8 (n=29)
3.6 (n=18)
2.8 (n=14)
2.6 (n=13)
2 (n=10)
2 (n=10)
1.8 (n=9)
1.8 (n=9)
1.6 (n=8)
1.6 (n=8)
1.6 (n=8)
1.6 (n=8)
1.4 (n=7)
1.4 (n=7)
1.2 (n=6)
1.2 (n=6)
1.2 (n=6)
1.2 (n=6)
1 (n=5)
1 (n=5)
1 (n=5)
0.8 (n=4)
0.8 (n=4)
0.8 (n=4)
0.8 (n=4)
0.8 (n=4)
0.8 (n=4)
0.6 (n=3)
0.6 (n=3)
0.6 (n=3)
0.6 (n=3)
0.6 (n=3)
0.6 (n=3)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.4 (n=2)
0.2 (n=1)
0.2 (n=1)
0.2 (n=1

Figure S2. Heatmap of the source proteins of the immunopeptidome of ECN90 β cells and primary human islets. A-B. Relative representation of source proteins in ECN90 β cells (A) and human islets (B), along with the corresponding significant log2 fold changes (FC) and percent values, ordered from the most represented protein in untreated cells. **C.** Relative representation of source proteins in two different primary islet preparations. In all panels, relative representation of each source protein is calculated based on the percent of unique ligands identified for each protein out of the total number of ligands in a given condition. Ligands included herein are conventional, post-translationally modified and mRNA splice peptides. Peptides carrying PTMs were counted only for PTMs defined as likely biological; they were otherwise counted as unmodified. mRNA splice variants were considered only when validated by spectral matching.

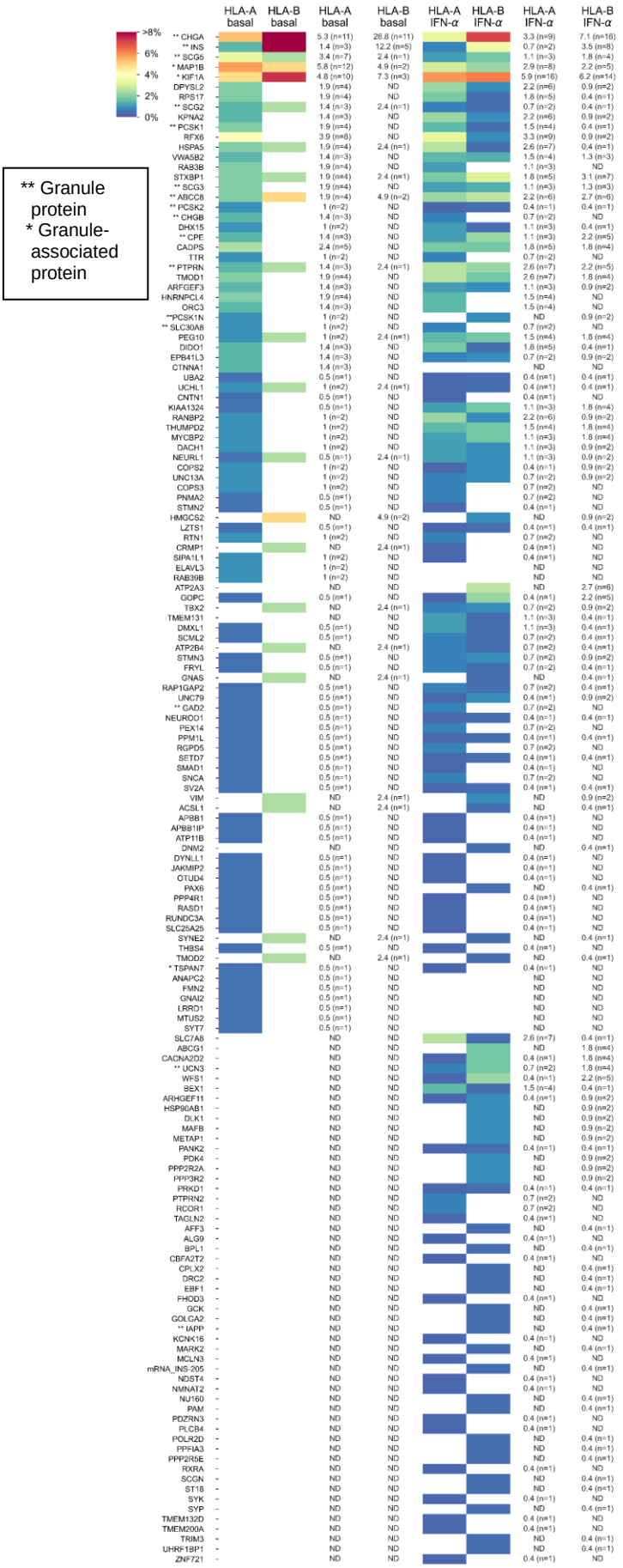


Figure S3. Heatmap of the source proteins of the immunopeptidome of ECN90 β cells separated by predicted HLA-I restriction. Legend is the same as in Fig. S2.

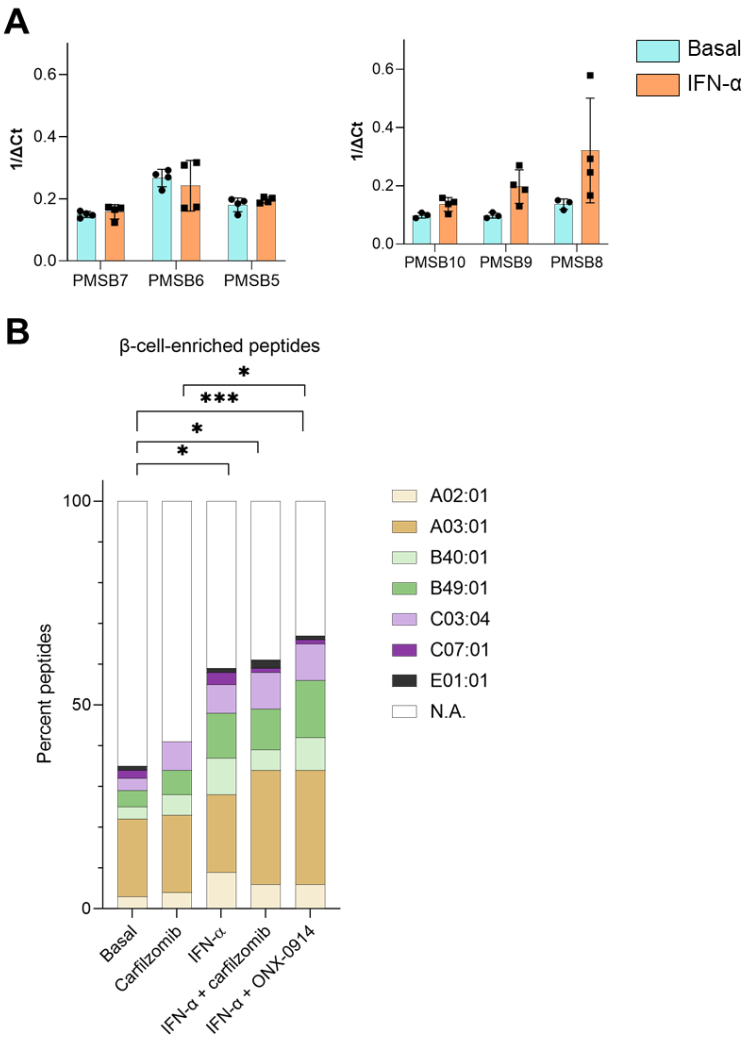


Figure S4. Inhibition of constitutive and immuno-proteasome on ECN90 β cells. A. Expression of the subunits of the constitutive (left) and immuo-proteasome (right) in ECN90 β cells assessed by qPCR (n=4 biological replicates). **B.** Relative distribution of predicted HLA-I ligands for each allele expressed by ECN90 β cells across conditions. *P<0.05, ***P<0.001 by χ^2 test.

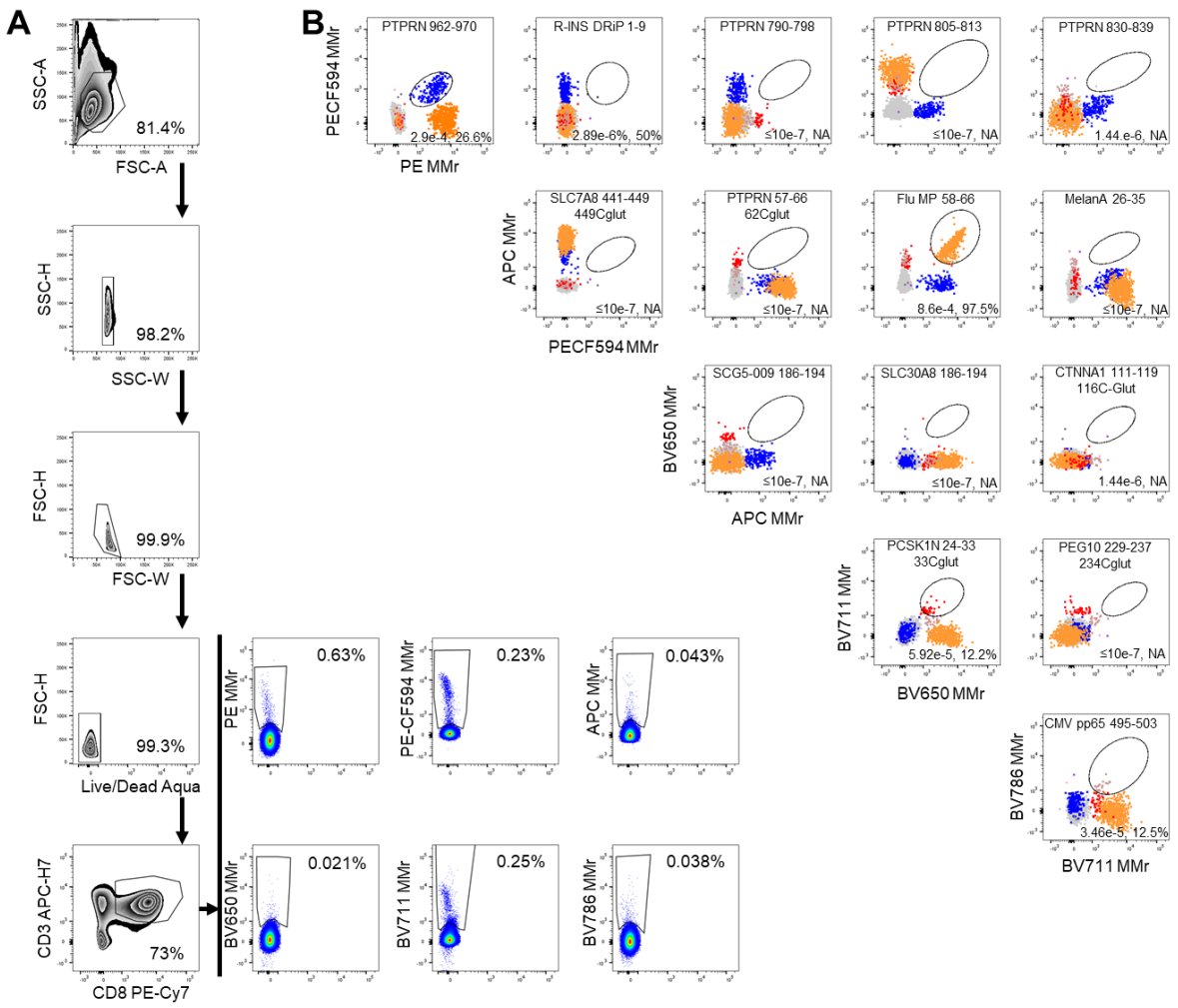


Figure S5. Gating strategy for the combinatorial analysis of HLA-A2-restricted islet-reactive MMr$^+$CD8$^+$ T cells. A. Frozen/thawed PBMCs from healthy donor H547C were magnetically depleted of CD8$^-$ cells prior to staining. Following flow cytometry acquisition, cells were sequentially gated on small lymphocytes, singlets, live cells (Live/Dead Aqua$^-$), CD3$^+$CD8$^+$ T cells, and total PE$^+$, PE-CF594$^+$, APC$^+$, BV650$^+$, BV711$^+$, and BV786$^+$ MMr$^+$ T cells. Using Boolean operators, these latter gates allowed selective visualization of each double-MMr$^+$ population by including only those events positive for the corresponding fluorochrome pair. **B.** The final readout obtained for the 15 peptides analyzed in one of the 5 staining panels acquired for each donor: 7 conventional islet peptides (including a literature-referenced (R-) peptide R-INS DRiP$_{1\text{-}9}$), 5 S-glutathionylated islet peptides, 3 control peptides (MelanA$_{26\text{-}35}$, CMV pp65$_{495\text{-}503}$ and Flu MP$_{58\text{-}66}$). Events corresponding to each epitope-reactive population are overlaid in different colors within each dot plot. Numbers in each panel indicate the MMr$^+$CD8$^+$ T-cell frequency out of total CD8$^+$ T cells and the percent effector/memory fraction among MMr$^+$ cells (NA when not assigned, i.e. <5 MMr$^+$ cells counted).

References

1. Hall, J. C. & Rosen, A. Type I interferons: crucial participants in disease amplification in autoimmunity. *Nat Rev Rheumatol* **6**, 40–49 (2010).

2. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat Rev Immunol* **14**, 36–49 (2014).

3. Taniguchi, T., Ogasawara, K., Takaoka, A. & Tanaka, N. IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol* **19**, 623–655 (2001).

4. Izaguirre, A. *et al.* Comparative analysis of IRF and IFN-alpha expression in human plasmacytoid and monocyte-derived dendritic cells. *J Leukoc Biol* **74**, 1125–1138 (2003).

5. Bon, A. L. *et al.* Type I Interferons Potently Enhance Humoral Immunity and Can Promote Isotype Switching by Stimulating Dendritic Cells In Vivo. *Immunity* **14**, 461–470 (2001).

6. Jego, G. *et al.* Plasmacytoid Dendritic Cells Induce Plasma Cell Differentiation through Type I Interferon and Interleukin 6. *Immunity* **19**, 225–234 (2003).

7. Luft, T. *et al.* Type I IFNs Enhance the Terminal Differentiation of Dendritic Cells. *J Immunol* **161**, 1947–1953 (1998).

8. Gallucci, S., Lolkema, M. & Matzinger, P. Natural adjuvants: Endogenous activators of dendritic cells. *Nat Med* **5**, 1249–1255 (1999).

9. Marroqui, L. *et al.* Interferon-α mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes. *Diabetologia* **60**, 656–667 (2017).

10. Marrack, P., Kappler, J. & Mitchell, T. Type I interferons keep activated T cells alive. *J Exp Med* **189**, 521–530 (1999).

11. Biron, C. A., Nguyen, K. B., Pien, G. C., Cousens, L. P. & Salazar-Mather, T. P. Natural killer cells in antiviral defense: function and regulation by innate cytokines. *Annu Rev Immunol* **17**, 189–220 (1999).

12. Kolumam, G. A., Thomas, S., Thompson, L. J., Sprent, J. & Murali-Krishna, K. Type I interferons act directly on CD8 T cells to allow clonal expansion and memory formation in response to viral infection. *J Exp Med* **202**, 637–650 (2005).

13. Curtsinger, J. M., Valenzuela, J. O., Agarwal, P., Lins, D. & Mescher, M. F. Type I IFNs provide a third signal to CD8 T cells to stimulate clonal expansion and differentiation. *J Immunol* **174**, 4465–4469 (2005).

14. Qaisar, N., Jurczyk, A. & Wang, J. P. Potential role of type I interferon in the pathogenic process leading to type 1 diabetes. *Curr Opin Endocrinol Diabetes Obes* **25**, 94–100 (2018).

15. Marroqui, L., Perez-Serna, A. A., Babiloni-Chust, I. & Dos Santos, R. S. Chapter One - Type I interferons as key players in pancreatic β-cell dysfunction in type 1 diabetes. in *International Review of Cell and Molecular Biology* (eds. Santin, I. & Galluzzi, L.) vol. 359 1–80 (Academic Press, 2021).

16. Stefan, M. *et al.* Genetic–epigenetic dysregulation of thymic TSH receptor gene expression triggers thyroid autoimmunity. *Proc Natl Acad Sci USA* **111**, 12562–12567 (2014).

17. Hasham, A. *et al.* Genetic analysis of interferon induced thyroiditis (IIT): evidence for a key role for MHC and apoptosis related genes and pathways. *J Autoimmun* **44**, 61–70 (2013).

18. Niewold, T. B. & Swedler, W. I. Systemic lupus erythematosus arising during interferon-alpha therapy for cryoglobulinemic vasculitis associated with hepatitis C. *Clin Rheumatol* **24**, 178–181 (2005).

19. Cacopardo, B., Benanti, F., Pinzone, M. R. & Nunnari, G. Rheumatoid arthritis following PEG-interferon-alfa-2a plus ribavirin treatment for chronic hepatitis C: a case report and review of the literature. *BMC Res Notes* **6**, 437 (2013).

20. Blanco, P., Palucka, A. K., Gill, M., Pascual, V. & Banchereau, J. Induction of Dendritic Cell Differentiation by IFN-α in Systemic Lupus Erythematosus. *Science* **294**, 1540–1543 (2001).

21. Ferreira, R. C. *et al.* A Type I Interferon Transcriptional Signature Precedes Autoimmunity in Children Genetically at Risk for Type 1 Diabetes. *Diabetes* **63**, 2538–2550 (2014).

22. Kallionpää, H. *et al.* Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility. *Diabetes* **63**, 2402–2414 (2014).

23. Chehadeh, W. *et al.* Increased Level of Interferon-α in Blood of Patients with Insulin-Dependent Diabetes Mellitus: Relationship with Coxsackievirus B Infection. *J Infect D* **181**, 1929–1939 (2000).

24. Lundberg, M., Krogvold, L., Kuric, E., Dahl-Jørgensen, K. & Skog, O. Expression of Interferon-Stimulated Genes in Insulitic Pancreatic Islets of Patients Recently Diagnosed With Type 1 Diabetes. *Diabetes* **65**, 3104–3110 (2016).

25. Foulis, A. K., Farquharson, M. A. & Meager, A. Immunoreactive alpha-interferon in insulin-secreting beta cells in type 1 diabetes mellitus. *Lancet* **2**, 1423–1427 (1987).

26. Huang, X. *et al.* Interferon Expression in the Pancreases of Patients With Type I Diabetes. *Diabetes* **44**, 658–664 (1995).

27. Dotta, F. *et al.* Coxsackie B4 virus infection of β cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. *Proc Natl Acad Sci USA* **104**, 5115–5120 (2007).

28. Nekoua, M. P., Alidjinou, E. K. & Hober, D. Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus. *Nat Rev Endocrinol* **18**, 503–516 (2022).

29. Richardson, S. J. *et al.* Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia* **59**, 2448–2458 (2016).

30. Colli, M. L. *et al.* An integrated multi-omics approach identifies the landscape of interferon-α-mediated responses of human pancreatic beta cells. *Nat Commun* **11**, 2584 (2020).

31. Javitt, A. *et al.* Pro-inflammatory Cytokines Alter the Immunopeptidome Landscape by Modulation of HLA-B Expression. *Front Immunol* **10**, 141 (2019).

32. Goncalves, G. *et al.* IFNγ Modulates the Immunopeptidome of Triple Negative Breast Cancer Cells by Enhancing and Diversifying Antigen Processing and Presentation. *Front Immunol* **12**, 645770 (2021).

33. Gonzalez-Duque, S. *et al.* Conventional and Neo-antigenic Peptides Presented by β Cells Are Targeted by Circulating Naïve CD8+ T Cells in Type 1 Diabetic and Healthy Donors. *Cell Metab* **28**, 946–960 (2018).

34. McGinty, J. W. *et al.* Recognition of Posttranslationally Modified GAD65 Epitopes in Subjects With Type 1 Diabetes. *Diabetes* **63**, 3033–3040 (2014).

35. Babon, J. A. B. *et al.* Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. *Nat Med* **22**, 1482–1487 (2016).

36. Aaron Wiles, T. *et al.* Identification of Hybrid Insulin Peptides (HIPs) in Mouse and Human Islets by Mass Spectrometry. *J Proteome Res* **18**, 814–825 (2019).

37. McCarthy, M. K. & Weinberg, J. B. The immunoproteasome and viral infection: a complex regulator of inflammation. *Front Microbiol* **6**, 21 (2015).

38. Lundh, M. *et al.* The immunoproteasome is induced by cytokines and regulates apoptosis in human islets. *J Endocrinol* **233**, 369–379 (2017).

39. Carré, A. & Mallone, R. Making Insulin and Staying Out of Autoimmune Trouble: The Beta-Cell Conundrum. *Front Immunol* **12**, 639682 (2021).

40. Wan, X. *et al.* Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* **560**, 107–111 (2018).

41. Liao, H. *et al. MARS: Improved De Novo Peptide Candidate Selection for Non-Canonical Antigen Target Discovery in Cancer*. (2022) doi:10.21203/rs.3.rs-1890352/v1, Preprint.

42. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res* **48**, W449–W454 (2020).

43. Kracht, M. J. L. *et al.* Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat. Med.* **23**, 501–507 (2017).

44. Da Silva Xavier, G. The Cells of the Islets of Langerhans. *J Clin Med* **7**, 54 (2018).

45. James, E. A., Pietropaolo, M. & Mamula, M. J. Immune Recognition of β-Cells: Neoepitopes as Key Players in the Loss of Tolerance. *Diabetes* **67**, 1035–1042 (2018).

46. Buitinga, M. *et al.* Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. *Diabetes* **67**, 2337–2348 (2018).

47. Sodré, F. M. C. *et al.* Peptidylarginine Deiminase Inhibition Prevents Diabetes Development in NOD Mice. *Diabetes* **70**, 516–528 (2021).

48. Azoury, M. E. *et al.* CD8+ T Cells Variably Recognize Native Versus Citrullinated GRP78 Epitopes in Type 1 Diabetes. *Diabetes* **70**, 2879–2891 (2021).

49. Wang, X. *et al.* Accurate Identification of Deamidation and Citrullination from Global Shotgun Proteomics Data Using a Dual-Search Delta Score Strategy. *J Proteome Res* **19**, 1863–1872 (2020).

50. Culina, S. *et al.* Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Sci Immunol* **3**, eaao4013 (2018).

51. Azoury, M. E. *et al.* Peptides Derived From Insulin Granule Proteins Are Targeted by CD8+ T Cells Across MHC Class I Restrictions in Humans and NOD Mice. *Diabetes* **69**, 2678–2690 (2020).

52. James, E. A., Mallone, R., Kent, S. C. & DiLorenzo, T. P. T-Cell Epitopes and Neo-epitopes in Type 1 Diabetes: A Comprehensive Update and Reappraisal. *Diabetes* **69**, 1311–1335 (2020).

53. Alanio, C., Lemaitre, F., Law, H. K. W., Hasan, M. & Albert, M. L. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. *Blood* **115**, 3718–3725 (2010).

54. Shimizu, F., Shimizu, M. & Kamiyama, K. Inhibitory Effect of Interferon on the Production of Insulin. *Endocrinology* **117**, 2081–2084 (1985).

55. Nordmann, T. M. *et al.* The Role of Inflammation in β-cell Dedifferentiation. *Sci Rep* **7**, 6285 (2017).

56. Lee, H. *et al.* Beta Cell Dedifferentiation Induced by IRE1α Deletion Prevents Type 1 Diabetes. *Cell Metab* **31**, 822-836.e5 (2020).

57. Mallone, R., Halliez, C., Rui, J. & Herold, K. C. The β-Cell in Type 1 Diabetes Pathogenesis: A Victim of Circumstances or an Instigator of Tragic Events? *Diabetes* **71**, 1603–1610 (2022).

58. Braud, V. M. *et al.* HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* **391**, 795–799 (1998).

59. Jiang, H. *et al.* HLA-E-restricted regulatory CD8(+) T cells are involved in development and control of human autoimmune type 1 diabetes. *J Clin Invest* **120**, 3641–3650 (2010).

60. Walters, L. C. *et al.* Pathogen-derived HLA-E bound epitopes reveal broad primary anchor pocket tolerability and conformationally malleable peptide binding. *Nat Commun* **9**, 3137 (2018).

61. Chong, C. *et al.* High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferonγ-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Mol Cell Proteomics* **17**, 533–548 (2018).

62. Girdlestone, J., Isamat, M., Gewert, D. & Milstein, C. Transcriptional regulation of HLA-A and -B: differential binding of members of the Rel and IRF families of transcription factors. *Proc Natl Acad Sci USA* **90**, 11568–11572 (1993).

63. Anderson, A. M. *et al.* Human islet T cells are highly reactive to preproinsulin in type 1 diabetes. *Proc Natl Acad Sci USA* **118**, e2107208118 (2021).

64. Samassa, F. & Mallone, R. Self-antigens, benign autoimmunity and type 1 diabetes: a beta-cell and T-cell perspective. *Curr Opin Endocrinol Diabetes Obes* **29**, 370–378 (2022).

65. Thomaidou, S. *et al.* Long RNA Sequencing and Ribosome Profiling of Inflamed β-Cells Reveal an Extensive Translatome Landscape. *Diabetes* **70**, 2299–2312 (2021).

66. Mann, S. E. *et al.* Multiplex T Cell Stimulation Assay Utilizing a T Cell Activation Reporter-Based Detection System. *Front Immunol* **11**, (2020).

67. Kim, S. K., Woodcroft, K. J., Khodadadeh, S. S. & Novak, R. F. Insulin Signaling Regulates γ-Glutamylcysteine Ligase Catalytic Subunit Expression in Primary Cultured Rat Hepatocytes. *J Pharmacol Exp Ther* **311**, 99–108 (2004).

68. Domínguez, C., Ruiz, E., Gussinye, M. & Carrascosa, A. Oxidative Stress at Onset and in Early Stages of Type 1 Diabetes in Children and Adolescents. *Diabetes Care* **21**, 1736–1742 (1998).

69. Darmaun, D. *et al.* Evidence for Accelerated Rates of Glutathione Utilization and Glutathione Depletion in Adolescents With Poorly Controlled Type 1 Diabetes. *Diabetes* **54**, 190–196 (2005).

70. Kohnert, K.-D., Jahr, H., Schmidt, S., Hahn, H.-J. & Zühlke, H. Demonstration of insulin degradation by thiol-protein disulfide oxidoreductase (glutathione-insulin transhydrogenase) and proteinases of pancreatic islets. *Biochimica et Biophysica Acta (BBA) - Enzymology* **422**, 254–259 (1976).

71. Trujillo, J. A. *et al.* The Cellular Redox Environment Alters Antigen Presentation. *J Biol Chem* **289**, 27979–27991 (2014).

72. Chen, W., Yewdell, J. W., Levine, R. L. & Bennink, J. R. Modification of Cysteine Residues In Vitro and In Vivo Affects the Immunogenicity and Antigenicity of Major Histocompatibility Complex Class I–restricted Viral Determinants. *J Exp Med* **189**, 1757–1764 (1999).

73. Pierce, R. A. *et al.* Cutting Edge: The HLA-A*0101-Restricted HY Minor Histocompatibility Antigen Originates from DFFRY and Contains a Cysteinylated Cysteine Residue as Identified by a Novel Mass Spectrometric Technique. *J Immunol* **163**, 6360–6364 (1999).

74. Kronenberg-Versteeg, D. *et al.* Molecular Pathways for Immune Recognition of Preproinsulin Signal Peptide in Type 1 Diabetes. *Diabetes* **67**, 687–696 (2018).

75. Rondas, D. *et al.* Citrullinated Glucose-Regulated Protein 78 Is an Autoantigen in Type 1 Diabetes. *Diabetes* **64**, 573–586 (2014).

76. Hebbandi Nanjundappa, R. *et al.* A Gut Microbial Mimic that Hijacks Diabetogenic Autoreactivity to Suppress Colitis. *Cell* **171**, 655-667.e17 (2017).

77. Tai, N. *et al.* Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. *J Exp Med* **213**, 2129–2146 (2016).

78. Lanz, T. V. *et al.* Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature* **603**, 321–327 (2022).

79. Girdhar, K. *et al.* A gut microbial peptide and molecular mimicry in the pathogenesis of type 1 diabetes. *Proc Natl Acad Sci USA* **119**, e2120028119 (2022).

80. Bentzen, A. K. *et al.* Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol* **34**, 1037–1045 (2016).

81. Colli, M. L. *et al.* PDL1 is expressed in the islets of people with type 1 diabetes and is up-regulated by interferons-α and-γ via IRF1 induction. *EBioMedicine* **36**, 367–375 (2018).

82. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).

83. Harrow, J. *et al.* GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res* **22**, 1760–1774 (2012).

84. Segerstolpe, Å. *et al.* Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell Metab* **24**, 593–607 (2016).

Discussion and perspectives

In this thesis, I have investigated the antigenic visibility of β cells exposed to the earliest suspected trigger of T1D (article #1) and characterized the effect of the subsequent inflammatory state (article #2). This work sheds light on some key mechanisms supporting the evidence that β cells may participate in their own destruction rather than being passive victims. Since a detailed discussion is already provided in the corresponding section of each article, I here chose to limit redundancy and focus on points not covered in the articles and the perspective they open for ongoing and future experiments.

Immune evasion of CVB-infected β cells

Using an *in vitro* model of CVB infection and an immunopeptidomics approach, we first characterized HLA-I-bound viral peptides presented by infected β cells. Only a few HLA-I-bound viral peptides were retrieved. These HLA-I-eluted peptides and *in silico* predicted binders that we identified as HLA-A*02:01 (A2) or -A*03:01 (A3) ligands were recognized by circulating $CD8^+$ T cells of CVB-seropositive donors to a limited extent (4/17-6/28 of HLA-eluted and *in silico* predicted peptides validated for T-cell recognition). Of note, most HLA-I eluted sequences were predicted to bind to the HLA-A3 allele, which may reflect a bias of MS. Indeed, due to its initial use for trypsin-digested samples in proteomics, MS works best with tryptic and semi-tryptic peptides, i.e. carrying a lysine or arginine residue at their C- and/or N-terminus. Given the semi-tryptic HLA-A3 binding motif, this MS feature makes the identification of HLA-A3 candidates easier.

Surprisingly, most of HLA-eluted CVB peptides mapped to non-structural proteins. The translation of the CVB genome leads to the production of a single polyprotein that needs to be cleaved to ensure the completion of CVB replication. Because it is always the first protein translated, we were expecting the structural protein P1 to be the most represented one in infected cells. However, only few peptides mapped to P1, and none of them were significantly recognized by $CD8^+$ T cells. Non-structural proteins may be a better substrate to feed the antigen processing and presentation pathway, as they are not encapsulated in newly-formed virions thus making them more abundantly and more easily available for degradation. Unfortunately, only a few studies examined the proteome of infected β cells, and none of them characterized the relative abundance of structural and non-structural viral proteins[387,388].

This weak presentation of HLA-bound CVB peptides illustrates the capacity of CVB to evade immune recognition and response. CVB immune escape mechanisms involve disrupting the host membrane trafficking[44], which leads to limited presentation of newly-formed pHLA-I

complexes as well as internalization of surface HLA molecules, explaining both the low number of CVB peptides retrieved and the reduced HLA-I expression observed. This disruption also leads to altered insulin secretion. Therefore, this disturbed vesicle trafficking likely drives changes in the entire immunopeptidome of the host cell. One of the limitations of our study is its focus on CVB epitopes presented by infected β cells, disregarding the possible modifications of the endogenous immunopeptidome and their impact on β-cell interactions with the immune system. However, on average, we identified 33-61x more 8-12mers in basal and 53-100x more 8-12mers in the IFN-α conditions, as compared to infected cells, while we have used only 1.5-4x and 1.2-3x more cells in the basal and in IFN-α-exposed cells, respectively (Figure 23). The use of different instruments for the two sets of experiment surely accounts for some differences, but more generally these numbers emphasize on the dramatic diminution of the immunopeptidome of infected ECN90 β cells. While we expected fewer peptides to be presented upon CVB infection, the amplitude of this variation also suggests drastic changes in the identity of displayed sequences.

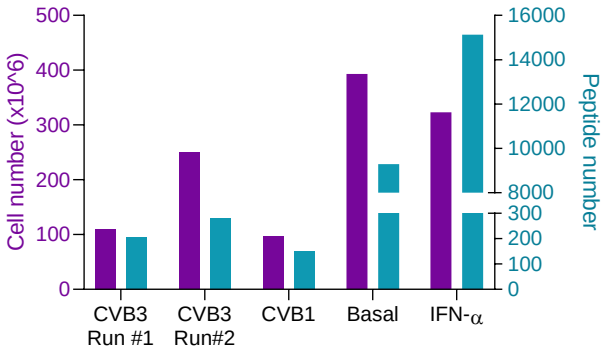


Figure 23. Cell and peptides numbers across experiments.

When analyzing $CD8^+$ T cells recognizing specific CVB epitopes in seropositive adults, no difference in frequency nor in phenotype could be observed. Indeed, their frequency (~5-50/10^6) was overall marginally higher than that reported for autoimmune $CD8^+$ T cells recognizing islet epitopes (1-50/10^6)[119,188]. Moreover, the effector/memory fraction within these antigen-reactive populations was limited for several individuals analyzed. In comparison, the recognition of the control peptide from Influenzavirus consistently led to high frequencies and mainly effector/memory phenotypes. Although the peak of incidence of CVB infections is during childhood, this observation did not reflect a long delay since exposure in adults, because the same observations were repeated in non-diabetic and T1D children. These features suggest

an overall poor $CD8^+$ T cell response. To support this interpretation, these anti-CVB T-cell responses need to be longitudinally characterized before and after CVB exposure. To this end, experiments are planned using stored longitudinal PBMC samples from the Finnish DIPP cohort. In this cohort, PBMC, serum and stool samples have been collected every 6 months starting at 6 months of age, with information about anti-CVB antibody seroconversion and detection of viral RNA in the stools available. This will allow to precisely track changes in $CD8^+$ T cells response before and after CVB encounter.

Another explanation for the results we obtained is that the peripheral blood does not necessarily mirror the local immune status, as T cells engaged in the anti-viral response may be sequestered in tissues or draining lymph nodes. Even though the frequency of CVB-reactive $CD8^+$ T cells is not increased in the spleen, their phenotype is mainly effector/memory (90-100% of the cells $CD45RA^+CCR7^-$, $CD45RA^-CCR7^+$ or $CD45RA^-CCR7^-$). For one donor, whose splenocytes showed high reactivity to $CVB1_{1356\text{-}1364}$/$CVB3_{1359\text{-}1367}$ peptide, available pLNs were clonally expanded and displayed reactivity to multiple CVB epitopes. Of interest, we could compare the blood, spleen and pLNs from non-diabetic and T1D donors, and we observed a trend towards higher frequencies of CVB-reactive T cells in pLNs of the few T1D donors available, with an overall exhausted ($PD\text{-}1^+$) phenotype, further emphasizing their tissue sequestration. When sequencing the TCR of CVB-reactive $CD8^+$ T cells, we detected the presence of repeated private clonotypes, thus highlighting previous *in vivo* expansion, which correlates with the effector/memory phenotype observed. The next step will be to search for the presence of these $CD8^+$ T cells by *in situ* MMr staining on pLN and pancreas tissue slides from nPOD specimens. This will inform us on the tissue localization of these CVB-reactive $CD8^+$ T cells. In light of the poor viral epitope presentation by infected b cells, we may expect to find CVB-reactive $CD8^+$ T cells in the pLNs rather than in the pancreas.

The finding of a poor anti-CVB immune response, combined with results gained from our cytotoxic assays investigating the T-cell and the CVB lysis capacity, strongly suggest that anti-viral $CD8^+$ T-cell responses are not the main drivers of the destruction of CVB infected β cells and the subsequent trigger of autoimmunity. This destruction is rather mediated by CVB itself, which lysed β cells *in vitro*. These results lend rationale to ongoing strategies aimed at vaccinating individuals against CVB to boost T-cell memory and prevent T1D onset. This is one of the extended application of the CVB pentavalent inactivated vaccine[55] currently tested in seronegative adult subjects (NCT04690426). Of note, this vaccine only contains structural proteins. Hence, using our MMr loaded with epitopes from structural or non-structural viral

proteins, we should be able to discriminate between T-cell responses to vaccination (i.e. against structural epitopes) and to natural infection (i.e. against both structural and non-structural epitopes). Thus, our study provides key tools to monitor vaccination effects at the T-cell level.

In this context of CVB vaccination strategy, our demonstration of a cross-recognition between a non-structural CVB peptide and a GAD homologous sequence deserves attention. As the vaccine only includes structural proteins, this cross-reactivity will not elicit any autoimmune off-target response. It will nonetheless be important to assess whether $CD8^+$ T cells recognizing this CVB/GAD-cross-reactive epitope can lyse non-infected β cells. This $GAD_{257\text{-}265}$ (KMFPEVKEK) homologous peptide was not retrieved at the surface of infected cells but was consistently found on resting and IFN-α exposed β cells and was one of the only granule-specific peptides whose abundance increased with IFN-α treatment (Figure 24). If cross-reactive $CD8^+$ T cells were able to lyse uninfected β cells, it would mean that T cells primed for the recognition of the viral peptide could wrongly kill neighboring uninfected cells displaying the GAD peptide, a phenomenon which might even be boosted by the inflammation resulting from the infection.

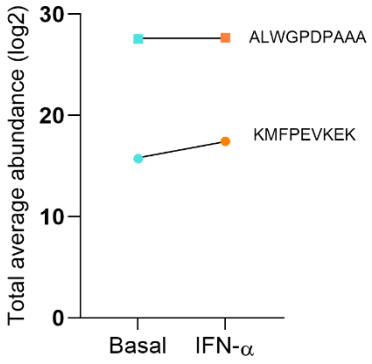


Figure 24. Abundance of $INS_{15\text{-}24}$ (ALWGPGPAAA) and $GAD_{257\text{-}265}$ (KMFPEVKEK) peptides
The peptides were identified by Peaks and quantified by Progenesis.

Our efforts contributed to characterize the viral epitopes presented by infected β cells and the corresponding $CD8^+$ T cell-response, that we found overall weak in both children and adults. From these findings, we now hypothesize that the link between CVB and T1D might rely on a faulty immune response upon entry of the virus in the organism, leading to its spreading to the pancreas. To challenge this hypothesis, we need to decipher the nature of the immune response at the entry site of the virus, namely in enterocytes. We will characterize the viral immunopeptidome of infected enterocytes and compare it to the one displayed by β cells,

to assess whether the viral immune escape mechanisms are as efficient at the entry site as they are in β cells. We will also estimate the contribution of the virus to cell lysis as opposed to T-cell mediated lysis. Ultimately, we will also determine the peptides presented by dendritic cells exposed to infected enterocytes or β cells and which might be priming T cells

Finally, the current study involves a high multiplicity of infection of β cells, mimicking an acute infection. However, it is more likely that repeated low titer infection plays a causal role in the onset of T1D[350]. Thus, our observations might not reflect the cascade of events leading to T1D. The poor $CD8^+$ T-cell response we observed with these settings suggests that an infection with a lower viral load would lead to an even weaker T-cell response. Early reports suggest that the contribution of $CD8^+$ T cells in CVB clearance is minimal and that their main participation is the production of cytokines rather than a direct cytolytic effect[389]. $CD8^+$ and $CD4^+$ T-cell-mediated immunity may in fact play a dual role depending on the viral titers, participating to viral clearance in high titers infections but being detrimental in low titers conditions[390]. Thus, as the immune response against CVB infection is limited in all individuals, T1D is more likely to arise from the consequences of viral infection, such as virus-mediated lysis of β cells and the resulting inflammation.

IFN-α increases the antigenic visibility of β cells

The effect of the inflammatory milieu that follows viral infection was the focus of the second part of my PhD project. Viral infection leads to the release of IFN-α by immune and non-immune-cells[294,295,391]. This cytokine also happens to be the earliest one observed in individuals at-risk of developing T1D[86,296], further supporting its importance.

Using the same *in vitro* β-cell model as before, we sequenced by MS the HLA-I-bound peptides presented by cells exposed or not to IFN-α. Expectedly, we observed that IFN-α increases the surface HLA-I expression and the number of ligands. Interestingly, IFN-α exposure also led to higher peptide abundance, suggesting that the inflammatory milieu expands the antigenic visibility of β cells both qualitatively and quantitatively. HLA-peptides in IFN-α exposed cells also originated from a more diversified set of proteins, likely reflecting the higher number of misfolded proteins that undergo proteasomal degradation due to the induction of ER stress. In parallel, fewer peptides originated from secretory granule proteins and contrary to all the other peptides, their abundance was not increased upon IFN-α exposure. One possible explanation resides in the reported inhibitory effect of IFNs on the production of insulin[362]. Another possibility is the previously-observed down-regulation of β-cell-specific genes upon

inflammation concomitantly to the upregulation of progenitor genes that are normally suppressed or expressed to a low level, leading to β-cell dedifferentiation[363]. Ultimately β-cell dedifferentiation alters their phenotype and their ability to secrete insulin and associated proteins. We will investigate both possibilities by measuring insulin content and secretion of cells exposed to IFN-α and by assessing the expression of β-cell specific and progenitor genes.

Yet, we have identified new antigens from the secretory granules (CHGB, SCG2, SCG3). All the antigens contained in secretory granules share the same features: they are soluble proteins that require enzymatic processing to form their bioactive products. The biosynthesis of secretory granules by β cells is highly adaptable and depends on the metabolic demands and stressors, which make their content prone to degradation and diversion toward the HLA-I antigen procession and presentation pathway[178]. It is worth to highlight here that, as reported in our previous work[188] and in spite of a higher depth of analysis (nearly 10-fold), some known β-cell antigens remain underrepresented (ZnT8, GAD, IAPP) or even absent (IGRP, INS DRiP) in the immunopeptidome of the β-cell line. This may be a bias of the model as for all of these proteins a few ligands were found in human islets, although the amount of starting material was much lower. These exceptions aside, the comparison of the immunopeptidome of HLA-A2$^+$/A3$^+$ human islets overlapped to a good extent to the one of our model cell line, emphasizing its relevance. Of note, the immunopeptidome of human islets contained more glucagon-derived peptides than insulin-derived ones suggesting that these cell types have higher antigenic visibility yet no autoimmune immunogenicity, the reasons of which are a focus of the laboratory for the next years.

One of the most striking effects of IFN-α on the β-cell immunopeptidome is the drastic increase in HLA-B ligands. There are two IFN-response elements upstream of the HLA-B locus versus only one for HLA-A and -C loci, leading to a higher HLA-B expression. Additionally, the source proteins of HLA-B ligands are much more limited than HLA-A and tend to be different from the ones of HLA-A upon inflammation. Taken together, these results highlight a remarkably skewed β-cell immunopeptidome upon IFN-α exposure and support the genetic risk associated with HLA-B39:06[28], while questioning the wide attention given to HLA-A2 ligands in T1D. This focus is mainly driven by technical constraints, as HLA-A*02:01 is one of the most frequent alleles in the Caucasian population (~45% of the population is HLA-A*02:01$^+$ vs. ~10-15% for the most common HLA-B alleles[392]), which makes it the easiest to study. Nevertheless, our data advocates for HLA-B to be the center of attention in future studies instead of HLA-A. This is further reinforced with the observation that IFN-γ, a later cytokine

observed in T1D, leads to an even higher increase of HLA-B gene expression (Figure 25). Dedicated efforts will be required to circumvent the current limitations for the study of HLA-B ligands (i.e. the recruitment of blood and tissue donors) but may unravel crucial key mechanisms of T1D progression and could possibly sustain the development of novel biomarkers.

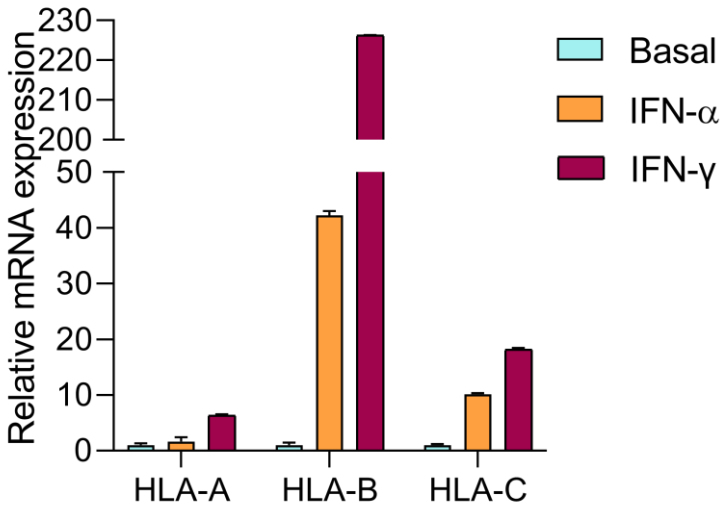


Figure 25. Expression of HLA genes upon cytokine exposure
Relative levels of mRNA expression of classical HLA-I alleles in ECN90 β cells exposed or not to IFN-α or to IFN-γ. GAPDH was used as an internal normalizing control and each gene was normalized to the basal sample. Data represent mean ± SEM of 3 biological replicates.

The presence of HLA-E ligands is another feature of the inflamed immunopeptidome. HLA-E binds leader sequences of other HLA-I molecules and engages the interaction with natural killer cells to ensure tolerance and self-surveillance[134]. In recent years, HLA-E has gained interest in T1D as it is upregulated specifically in ICIs[133] and preferentially in non-infiltrated ones[348] suggesting a role in the β-cell protection. Recent investigations have shown that HLA-E presents a limited set of peptides to regulatory $CD8^+$ T cells in T1D[365] and also displays pathogen-derived epitopes to $CD8^+$ T cells[136]. Besides the membrane-bound isoform, HLA-E also exists as an interferon-inducible soluble form, implying that it might be involved in bystander cell protection[133].

A key asset of our study also resides in the efforts put toward the identification of β-cell neo-epitopes that may drive T-cell responses. First, we used RNAseq data from human islets to select sequences giving rise to new protein isoforms and appended them to the search database. We retrieved 4 mRNA variants in HLA-eluted peptides from which 1 INS B-restricted mRNA candidate was confirmed. Second, we used a newly-developed unbiased algorithm to find

possible *cis*-spliced candidates[243]. Few peptides were determined as originating from two fragments of the same protein, but due to technical limitations in the discrimination of isoleucine from leucine residues, we could not pinpoint the exact source proteins and fragments. Therefore, we will proceed with proteasomal digestion of long precursor of β-cell-enriched proteins to narrow down the possibilities and to verify that the generation of the *cis*-splice candidates depends on proteasomal activity. Of note, most but not all of the *cis*-spliced peptides were specific to cells exposed to IFN-α. The overall contribution of proteasome-spliced peptides to disease immunogenicity seems however debated and low[393], thus we will need to confirm the pertinence of our candidates in T1D development by documenting the existence, frequency and phenotype of reactive $CD8^+$ T cells in diabetic and healthy donors. Third, we screened our data for relevant, possibly biological PTMs. We have confirmed the site of 12 PTMs basing our approach on spectral matching between modified and unmodified synthetic peptides and original spectral identification. Unless using strategies for the identification of specific PTMs, which usually require dedicated pre-MS sample preparation, the identification of biological PTMs by MS remains challenging. However, the most represented PTM in our data, S-glutathionylation, is unlikely to be an artifact due to its size and corresponds to the formation of a disulfide bond between a cysteine residue and a glutathione tripeptide. Glutathione is one of the major antioxidants in cells, its synthesis is promoted by insulin and its levels are reduced in T1D patients[373,374]. While the frequency of this modification may reflect the redox balance or imbalance of the cells, it can alter T-cell recognition[376] making glutathionylated peptides attractive neo-epitope candidates. Despite being thorough, our analysis is not exhaustive and several types of neo-epitopes recently reported in cancer, were not investigated and would require specialized analyses[240].

Lastly, as several drugs targeting the antigenic processing and presentation pathway are in clinical development and use[394,395], making it an easily targetable component, we have investigated the role of the constitutive and immuno-proteasome in shaping the immunopeptidome of β cells by inhibiting the chymotrypsin activity of both complexes. Clear patterns of presentation were observed, and we were able to group peptides based on their dependence on both the expression of these enzymatic complexes and the exposure to IFN-α. One HLA-ligand from the signal peptide of insulin was consistently observed, suggesting that its generation does not rely on the catalytic activity of the proteasome and of its inflammation-induced counterpart. This observation corroborates reports describing the peculiar pathway of presentation that fragments of the signal peptides can engage in[379].

The next step of this project will focus on determining the peptides that are recognized by $CD8^+$ T cells. For that purpose, we aim at using combinatorial tetramers for HLA-A2 and -A3 candidates while we will use DNA-barcoded tetramers for HLA-B restricted ligands. Our first screening of HLA-A2-ligands on PBMCs from 5 adult healthy donors indicates that the identified peptides are recognized by T cells that are largely naïve. These results will help us to select the candidates to investigate further in age- and sex-matched adult and pediatric cohorts of non-diabetic, at-risk and T1D donors. We will also assess the recognition of modified sequences, especially the S-glutathionylated peptides. Furthermore, despite identifying the peptides presented at the surface of β cells, we currently lack evidence of their display by antigen-presenting cells. This will be investigated to confirm the nature of the antigens that prime $CD4^+$ and $CD8^+$ T cells. Of interest, the secretory function of β cells points to a possibility for granule contents, released in the islet microenvironment and in the circulation, to be uptaken, processed and presented at distance by antigen-presenting cells [313]. To gain insight on the immune dialogue, which might initiate autoimmune mechanisms, we will expose monocyte-derived dendritic cells from donors carrying the risk allele HLA-DR4, to the secretome of β cells and characterize their HLA-II immunopeptidome.

Technical limitations and improvements

One of the general limitations of immunopeptidomics is the requirement for a high amount of starting material, which is not always possible. This issue was partly circumvented for small samples as this study benefitted from the latest MS technological advances with the use of trapped ion mobility spectrometry (tims) coupled to high-resolution time-of-flight (TOF) MS. The addition of this fourth dimension on the MS acquisition allows for a deeper sequencing ability increasing drastically the number of ligands identified. This innovation poses the challenge of peptide identification, as software needs to handle a new type of data. Peaks X Pro has a novel dedicated module that allows it to manage such data. However, Peaks may currently not be the best performing algorithm for the analysis of timsTOF data[396]. In our hands, using MSFragger[397] indeed increased the number of identifications (data not shown), but as the search algorithm can impact the repertoire of peptides[398], we decided to keep using Peaks for consistency across the study. TimsTOF datasets are much larger, which prevented us from using the Peaks PTM search and thus also limited the identification of PTMs to a few selected ones. The difference between algorithms with timsTOF data may require further investigation to ensure maximal recovery of relevant sequences.

Importantly, our work focuses on predicted HLA-I ligands, but the algorithm used to compute the predictions is trained with available datasets, thus imposing a bias toward the biology of cancer cells, which may be less accurate for β cells. Moreover, non-binders were completely overlooked, even though they have recently been put under the spotlight in the cancer field, with reports showing their importance as immunogenic targets[399,400]. The binding capacity is a good predictor of immunogenicity in the case of viral peptides, but it seems less critical for tumor and self-epitopes[254,399]. Therefore, it is possible that weak binding yet immunogenic β-cell peptides are discarded with our approach.

One of the strengths of our work is the implementation of an automated pipeline for data analysis, allowing easy filtering and sorting of peptides of interest across experiments. However, the "β-cell enriched" filter is biased toward data contained in previously published RNAseq datasets on human islets[386]. In light of a recent proteomics investigation of human islets[401], the "β-cell enriched" filter may need to be redefined to incorporate these new findings in our future investigations.

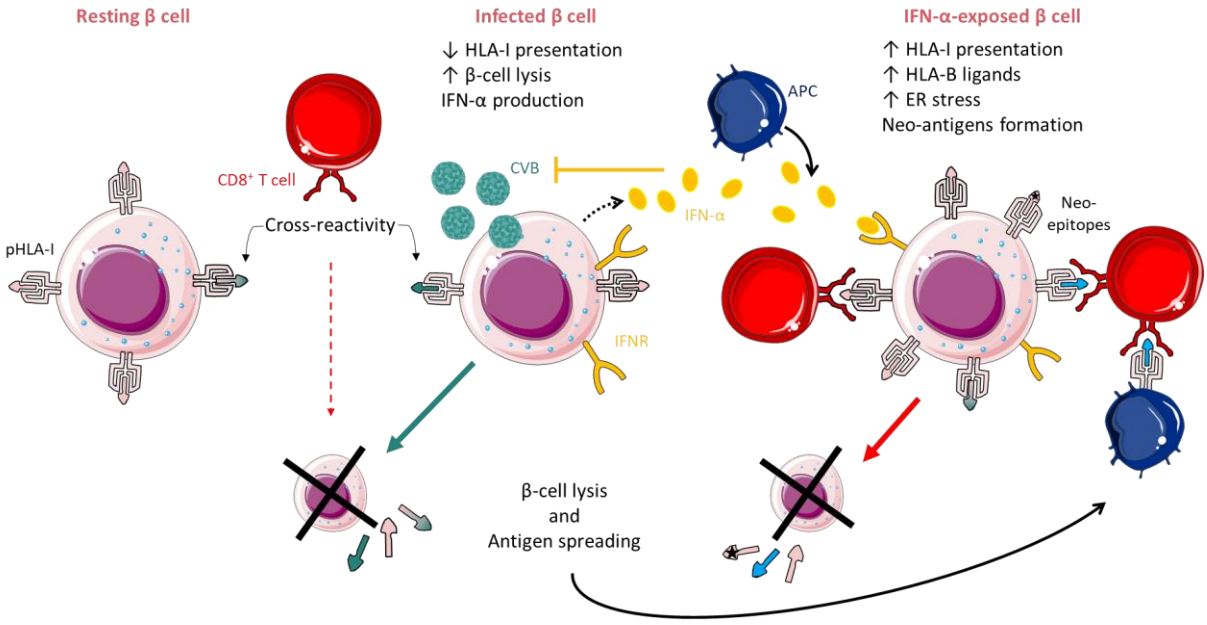


Figure 26. CVB- and IFN-α-mediated mechanisms participating in triggering β-cell autoimmunity

CVB-infected β cells display fewer HLA molecules containing a restricted set of viral and endogenous peptides. Once CVB completes its replication cycle, the β cell is lysed inducing antigen spreading. $CD8^+$ T cells recognizing viral epitopes possibly cross-react with β-cell antigens presented to a higher extent by uninfected cells, which may then undergo T-cell-mediated lysis. The viral infection also leads to IFN-α secretion by immune, and probably also by non-immune cells, to block virus spreading. Upon binding with its receptor, IFN-α prompts ER stress and the subsequent formation of neo-antigens. Together with the enhanced HLA-I expression, this leads to an increased antigenic visibility and raises the odds that $CD8^+$ T cells recognize their cognate target at the β-cell surface and lyse them, further fueling the autoimmune process through the release of antigens. These peptides may be taken up, processed, and presented by antigen-presenting cells to prime autoreactive T cells.

As a whole, this project unravels the antigenic visibility of β cells in light of the most credible environmental trigger of T1D and of its subsequent inflammation. The data collected clearly demonstrate that CVB hijacks the host cellular machinery to prevent its detection by the immune system, as revealed by a poor $CD8^+$ T-cell response. As a consequence, the endogenous β-cell peptides are also drastically less presented. On the contrary, IFN-α skews the immunopeptidome of β cells, boosting the diversity and quantity of HLA-I ligands including neo-epitopes, but also of presenting HLA-I molecules. Thus, since IFN-α expression can result from CVB infection, β cells which are under an immune ignorance state may become drastically more visible and vulnerable to autoimmune T cells (Figure 26). This wide fluctuation itself is likely to play a role in the switch from benign autoimmunity to T1D. Altogether, this project has many implications for the development of biomarkers, the monitoring of CVB vaccination and to further understand the early mechanisms of autoimmunity in T1D.

Annexes

Annex 1: Review about the diabetogenic mechanisms of CVB (invited article for *Endocrine Reviews*)

From association to causality: possible diabetogenic mechanisms of Coxsackievirus B infection and implications for type 1 diabetes prevention

Alexia Carré[1], Federica Vecchio[1], Malin Flodstrom-Tullberg[2], Sylvaine You[1], Roberto Mallone[1,3].

[1]Université Paris Cité, Institut Cochin, CNRS, INSERM, 75014 Paris, France.

[2]Karolinska Institutet, Department of Medicine Huddinge and Karolinska University Hospital, 14152 Stockholm, Sweden.

[3]Assistance Publique Hôpitaux de Paris, Hôpitaux Universitaires de Paris Centre-Université de Paris, Cochin Hospital, Service de Diabétologie et Immunologie Clinique, 75014 Paris, France.

Abstract

The evidence for an association between Coxsackievirus B (CVB) infection, pancreatic islet autoimmunity and clinical type 1 diabetes is increasing. Results from prospective cohorts and pancreas histopathology studies have provided a compelling case. However, the demonstration of a causal relationship is missing, and is likely to remain elusive until tested in humans by avoiding exposure to this candidate viral trigger. To this end, CVB vaccines have been developed and are entering clinical trials. However, the progress made in understanding the biology of the virus and in providing tools to address the long-standing question of causality contrasts with the scarcity of information about the antiviral immune responses triggered by infection. Beta-cell destruction may be primarily induced by CVB itself, possibly in the context of poor immune protection, or secondarily provoked by T-cell responses against CVB-infected beta cells. The possible participation of epitope mimicry mechanisms skewing the physiological antiviral response toward autoimmunity has also been suggested. We here review the available evidence for each of these three scenarios. Understanding which one is at play is critical to maximize the odds of success of CVB vaccination, and to develop suitable tools to monitor the efficacy of immunization and its intermingling with autoimmune onset or prevention.

The most outstanding question in the field of type 1 diabetes (T1D) concerns the identification of environmental triggers (1), a conundrum shared with many others autoimmune diseases. These environmental triggers are gaining importance, as the steady progression of T1D incidence in the last 50 years cannot be accounted by a genetic drift of the human population. Rather, the relative weight of predisposing gene variants (mostly mapping to HLA Class II loci) is decreasing. Indeed, T1D patients diagnosed in more recent years carry HLA haplotypes considered as protective to a larger extent than patients diagnosed decades ago, while high-risk haplotypes are becoming less common (2). This observation indirectly suggests that environmental pressure is increasing. Moreover, these environmental factors exert their role very early in life, as the majority (~64%) of children that later develop T1D display their first seroconversion for islet auto-antibodies (aAbs) during the first 2 years of life (3, 4).

A universal environmental trigger underlying all T1D cases is unlikely to exist, as disease heterogeneity, mostly related to age, is increasingly recognized (5). Nonetheless, the strongest evidence for an association with T1D points to Enterovirus infection, particularly to Coxsackieviruses B (CVBs) (6, 7). This echoes the robust evidence linking Epstein-Barr virus infection with multiple sclerosis (8-11). We will discuss current evidence for the association between CVB and T1D, and the possible pathogenic mechanisms at play. We contend that a better knowledge of such mechanisms is essential to inform the design of T1D prevention trials based on CVB vaccination (12, 13).

Common features of putative diabetogenic viruses

Many different viruses have been implicated as candidate triggers for T1D. These include Enterovirus (both CVB and Echovirus), Parechovirus and Rotavirus. The strength of evidence for each of those has been recently reviewed (14), but it is important to note that they share four key features:

1) They are responsible for infections during the first years of life, which may be in line with the early timing of aAb seroconversion in most pediatric T1D progressors (3, 4). CVB infections show some seasonal trends, being more common during the summer and early fall.

2) These infections are highly prevalent, with >95% of the general population being seropositive. It thus needs to be explained how such prevalent infections may trigger T1D in only few individuals.

3) They are transmitted mainly through the oro-fecal route. This is quite important, as pancreatic lymph nodes drain not only the pancreas itself but also parts of the intestinal tract, thus providing an ideal cross-road for immune cells and for CVB to reach the pancreas.

4) They can infect beta cells. They can thus exert their pathogenic mechanisms on the same cells that are targeted by T1D autoimmunity.

CVBs are classified into 6 serotypes (1 to 6) and belong to the family Picornaviridae and the genus Enterovirus, which also includes Poliovirus and Echovirus. They are small non-enveloped single-stranded RNA viruses that initially replicate in the submucosal lymph tissue of the distal small bowel, and of the upper respiratory tract when transmitted by aerosols. Although they usually cause asymptomatic infections, further dissemination to target organs can occur following a secondary viremia. This can occasionally result in severe diseases, such as meningitis, encephalitis, myocarditis, and systemic neonatal infections. Some of these clinical conditions, e.g. myocarditis, exemplify the capacity of CVBs to cause persistent (chronic) infections (7, 15-17).

Association between CVB infection and T1D: prospective and histopathology studies

The association between CVB infection and T1D is supported by temporal correlations from prospective cohorts and by spatial correlations documented in histopathological studies.

Temporal correlations. A 2011 meta-analysis of 26 case-control studies reported a significant association between Enterovirus infection and islet autoimmunity (odds ratio, OR 3.7) or clinical T1D (OR 9.8) (18), although some publication bias in favor of positive results is likely. The CVB serotypes that have been more frequently associated with T1D are CVB1 (19) and CVB4 (20), while some others, e.g. CVB3 and CVB6, have occasionally been suggested to be protective (19). Associations between CVB exposure (i.e. detection of CVB RNA in stools) and seroconversion and T1D progression were provided by prospective cohorts such as the Finnish T1D Prediction and Prevention (DIPP) (https://dipp.fi) and The Environmental Determinants of Diabetes in the Young (TEDDY; https://teddy.epi.usf.edu) studies. Based on serological evidence, i.e. neutralizing antibody (Ab) titers against different CVB serotypes, the group of H. Hyoty (19) reported that the prevalence of anti-CVB1 Abs was higher in children later seroconverting for islet aAbs, while that of anti-CVB3 and anti-CVB6 Abs was lower. When considered altogether, a positive CVB1 serology and a negative CVB3/CVB6 serology

conferred a 2.5 OR for subsequent islet aAb seroconversion. A protective role for maternal Abs was further suggested, as a negative CVB1 serology in cord blood and subsequent CVB1 seroconversion by 18 months of age was associated with a similar OR of 2.6. Given its focus on preclinical stage 1 (i.e. asymptomatic islet autoimmunity with normal insulin secretion), this study suggested that CVB exposure may be associated with aAb seroconversion rather than clinical progression. This is in line with previous studies from Cuba reporting an association between epidemic Echovirus infection and aAb seroconversion in the convalescent phase (21-23). This possibility was confirmed by a recent TEDDY study based on the longitudinal analysis of the fecal viral metagenome (20). The risk of future aAb seroconversion was not associated with short and independent infections, but rather with prolonged infections with the same Enterovirus B (mostly CVB4) serotype, with a risk increase of ~20% at each positive stool sample.

Spatial correlations. A spatial correlation between CVB infection and insulitis has been highlighted by immunohistochemistry (24-27), reverse-transcription polymerase chain reaction (27) and *in-situ* hybridization (28) using tissue specimens available through the network for Pancreatic Organ Donors with T1D (nPOD; https://www.jdrfnpod.org;), from brain-dead organ donors; the UK Exeter Archival Diabetes Biobank (EADB; https://foulis.vub.ac.be), from autopsy material; and the Norwegian Diabetes Virus Detection (DiViD) study, from living donors. Immunohistochemistry for the Enterovirus viral protein (VP)1 revealed a co-localization with T-cell infiltration and/or HLA Class I hyper-expression, which are the two hallmarks of insulitis. When considering these immunohistochemistry studies altogether (24-27), VP1+ islets were detected in ~70% of T1D cases and 14% of non-diabetic controls. All these studies also conclude that only a fraction (7-29%) of islets with residual insulin-containing cells are VP1+, suggesting that persistent low-grade rather than short-lived high-grade infections may occur, probably long before T1D clinical onset (12).

Possible mechanisms of persistent CVB infection

As discussed above, the small fraction of VP1+ cells that are present in the pancreas of T1D patients (corresponding to the late stage 3 disease in natural history) together with the temporal association between prolonged fecal CVB shedding and aAb seroconversion (i.e. early stage 1 disease) suggest a scenario of prolonged CVB infections. On the other hand, the few cases of aAb+ organ donors analyzed to date did not show evidence of acute, extensively lytic CVB

infection (5), although also in this case we only have a 'snapshot' available, which may date well beyond the time of CVB encounter. It should be noted that this does not exclude the possibility that a more limited lysis of CVB-infected beta cells may occur. The TEDDY stool virome study (20) further indicates that these prolonged CVB infections are the result of viral persistence (i.e. chronic infection following a single viral encounter) rather than recurrence (i.e. multiple viral encounters).

The viral mechanisms leading to persistent CVB infections have been recently reviewed (7). A first mechanism involves repeated CVB seeding from infection reservoirs in the gut (mostly duodenal epithelial cells), pancreas (ductal and beta cells) and blood cells (e.g. monocytes), through infections that can be either cytolytic (with release of new virions) or not (with virions released through other mechanisms, e.g. extracellular vesicles) (29). Reservoirs of viral persistence are also established in microvascular endothelial cells, without causing overt cytopathic effects but inducing the upregulation of adhesion molecules that may contribute to leukocyte recruitment (30). A second mechanism exploits a naturally occurring deletion in the 5′ non-coding region of the viral genome, which leads to reduced viral replication and persistent low-grade infection (31). Immune mechanisms may also participate, and include exacerbated innate interferon (IFN)-α responses (another hallmark of T1D, starting from its early preclinical stages) (32, 33) limiting viral replication and beta-cell lysis (34); and insufficient adaptive anti-viral responses that do not clear the virus (discussed below).

From association to causality: possible diabetogenic mechanisms of CVB infection

With these points in mind, three non-mutually exclusive mechanisms can be proposed to explain the triggering effect of CVB infection on islet autoimmunity (**Figure 1**):

1) *Direct* lytic effect of *CVB* on infected beta cells. This may imply *poor* adaptive immune responses unable to efficiently clear the virus.

2) *Indirect* lytic effect of *anti-viral immune responses* on infected beta cells, involving cytotoxic CD8+ T cells. This may imply *strong* adaptive immune responses leading to viral clearance.

Both mechanisms would lead to the release of beta-cell antigens in a pro-inflammatory environment, which may subsequently trigger islet autoimmunity.

3) Epitope mimicry. The physiological antiviral immune response may turn into a pathological autoimmune response against beta cells due to T-cell cross-reactivity between homologous CVB and beta-cell epitopes.

In the following sections, we will review the level of evidence for each of these mechanisms.

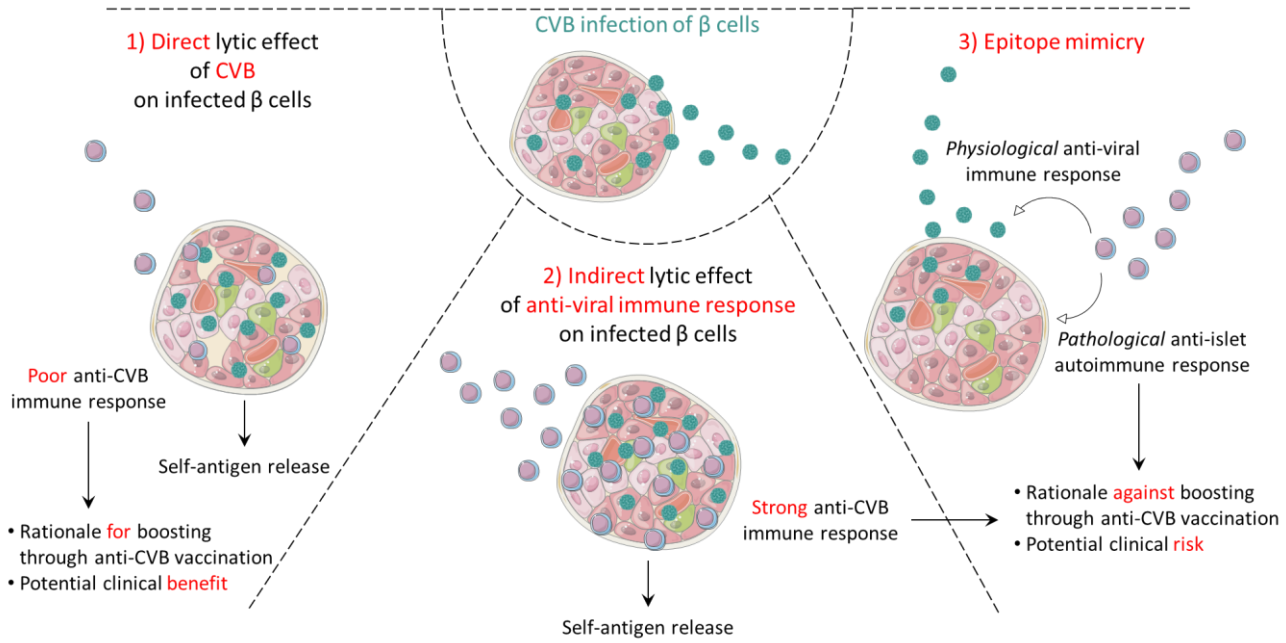


Figure 1. Three hypothetical, non-mutually exclusive mechanisms leading to β-cell autoimmunity, and implications for T1D prevention by means of CVB vaccination.

Direct lytic effect of CVB on infected beta cells.

CVB enters cells primarily by binding to surface Coxsackievirus and Adenovirus Receptor (CAR). Both human and murine alpha and beta cells express CAR and are CVB-permissive (34).

Mouse studies. Several *in-vivo* experimental models document the lytic effect of CVB on beta cells and its capacity to trigger diabetes without engaging potent anti-viral T-cell responses. These effects are dependent on the CVB serotype and the genetic background of the murine host. For instance, infection of C57BL/6 mice with CVB3/4 does not induce diabetes, while it does in 25% of infected SJL and CD1 hosts, independently of T cells (35, 36). This variable susceptibility is reminiscent of that observed with experimental CVB3-induced myocarditis (37). Most strains develop a severe acute myocarditis, but completely recover; few others develop chronic myocarditis associated with anti-cardiac myosin heavy chain aAbs that are also found in humans. This disease can be recapitulated by immunizing mice with cardiac myosin.

A direct lytic effect is also suggested in the multiple low-dose streptozotocin (LD-STZ) model, which triggers autoimmunity by inducing a moderate beta-cell lysis that releases autoantigens in an inflammatory environment. While a single LD-STZ injection was not diabetogenic, its combination with CVB3/4/5 inoculation induced hyperglycemia in CD1 mice (38, 39). Moreover, only a peri-islet insulitis pattern was observed (39), suggesting an accessory role for T cells in mediating beta-cell destruction.

Humanized immunodeficient mouse models further highlighted the beta-cell-lytic effect of CVB infection. These models are based on NOD/*scid*/*gamma* (NSG) mice deprived of endogenous islets by high-dose STZ or diphtheria toxin (DT) treatment (via a rat insulin promoter/DT-receptor transgene) and grafted with human islets (40, 41). Subsequent CVB4 infection led to hyperglycemia in 50% of animals 28 days later. Hyperglycemia was associated with CVB4 RNA and protein persistence in transplanted islets, increased endoplasmic reticulum (ER) stress and a type I IFN gene signature. Notably, pancreas histopathology revealed reduced insulin content but no significant islet destruction.

In spontaneously diabetes-prone non-obese diabetic (NOD) mice, CVB1/CVB3/CVB4 inoculation precipitates diabetes only when applied at the pre-diabetic stage - starting at 9 weeks of age, when invasive insulitis is already present (42-45). At this stage, CVB replicates efficiently in islets (46). On the contrary, CVB3 or CVB4 inoculation of young (<8-week-old) NOD mice with minimal insulitis has a preventative effect (42, 43), which is associated with a defective ability of CVB to replicate in islets, despite the detection of high viral titers in islets 1-2 days after inoculation and the expression of the CVB receptor CAR (43, 47). This CVB failure to replicate in islets has been associated with the induction of β-cell-intrinsic IFN responses and HIF-1α expression. Indeed, NOD mice with a β-cell-specific HIF-1α deficiency display accelerated diabetes upon CVB1/CVB4 infection, accompanied by increased pancreatic viral loads (48). This provides a mechanistic link between the susceptibility or resistance of beta cells to CVB infection and their capacity to sense viral RNA and induce intracellular anti-viral, IFN-mediated defense mechanisms. Accordingly, the inhibition of IFN responses with a beta-cell-specific transgene for the suppressor of cytokine signaling-1 (*socs-1*) rapidly induces diabetes after CVB1/CVB3/CVB4 infection, with very few residual insulin-positive cells (49-51). Also in this case, disease onset does not require T cells, as the outcome is similar in immunodeficient *socs-1*-transgenic NOD/*scid* mice (50).

The diabetes protection in young NOD mice versus acceleration at an older age seem at variance with the notion that, in humans, CVB infection may be an early trigger that precedes islet

autoimmunity (19, 20). Possible explanations include the fact that the persistent CVB infections that may release a critical self-antigen load in humans are difficult to reproduce in NOD mice, and that the magnitude and/or quality of anti-CVB immune responses may be different. In NOD mice, the requirement for persistent CVB infection may be bypassed by the critical threshold of beta-cell destruction already achieved by autoimmune mechanisms on their own. The relatively high CVB doses used in these infection models, which are administered through the intraperitoneal rather than the natural oral route, are another confounder.

Human studies. *Lytic effect on beta cells.* An early study assessed the lytic effect of several CVB and A strains on islet cells (52). The infected outer islet cells died over a few days and detached from islets, leading to smaller structures. No features of apoptosis were observed, rather suggesting necrosis with early chromatin condensation (pyknosis). Nonetheless, some beta cells in close proximity to infected and damaged beta cells remained virtually intact. This underlines the heterogeneous outcome of CVB infection (52), possibly reflecting its preferential replication in dividing cells and latent infection of quiescent cells (53), a feature shared by many viruses (54, 55).

Other effects on beta cells. CVB infection impacts beta cells on many other levels. First, the double-stranded (ds)RNA replicative intermediate is recognized by pattern recognition receptors such as toll-like receptor 3 (TLR3), retinoic-acid-inducible gene I (RIG-I) and melanoma differentiation associated protein 5 (MDA5) that trigger the production of pro-inflammatory cytokines, notably type I IFNs. Indeed, infection of human islets with CVB3, CVB4 or CVB5 induces type I IFN (mostly IFN-β) expression (34, 56), suggesting that the early blood type I IFN signature of T1D (32, 33) may reflect an anti-viral response. While this IFN response limits CVB replication and spreading, it also enhances beta-cell apoptosis (34). Additional effects of dsRNA on beta cells have been observed in an *in vitro* model using the synthetic dsRNA mimic polyinosinic-polycytidyilic acid (polyI:C) on a beta-cell line and primary human islets (57). PolyI:C downregulated beta-cell-specific genes (e.g. *INS, G6PC2, SLC30A8, MAFA*), induced *de-novo* expression of the progenitor-like transcription factor *SOX9* and impaired glucose-stimulated insulin secretion. This gene expression pattern suggests a dedifferentiation process and was recapitulated upon CVB5 infection. This dedifferentiated phenotype seems beta-cell-specific, as glucagon mRNA levels were unaffected in infected human islets (58). These polyI:C effects were triggered by the NF-κB and IFN regulatory factor pathways, and by the secretion of their downstream cytokines IFN-α and tumor necrosis factor (TNF)-α. Indeed, the use of IFN-α, alone or in combination with TNF-α, led to a similar *SOX9*

expression (57), suggesting that IFN-α may trigger a vicious cycle by disrupting the identity of neighboring cells in a paracrine fashion.

Other outcomes of CVB infection on different cells include intensive viral protein production, inhibition of host cell protein translation, impaired cellular calcium homeostasis and ER membrane modifications (59-61). Altogether, these alterations potentiate ER stress, therefore activating the unfolded protein response (UPR). The UPR is a natural cellular response to stress which aims at decreasing the translation rate, increasing the biosynthesis of protein-folding chaperones, and inducing the degradation of misfolded proteins. If the UPR fails to resolve cellular stress, it triggers apoptosis. Given their high insulin synthesis, beta cells have high basal levels of ER stress and naturally adjust the UPR to survive (62). This situation is easily decompensated by triggers such as CVB infection (63). Of further note, the anterograde vesicular trafficking is reduced, concomitantly with increased retrograde trafficking (64). In the highly secretory beta cells, this translates into reduced secretory activity and insulin stores (65), likely imposing an additional stress. The disruption of protein trafficking induced by CVB infection extends to autophagy, which is vital for beta cells to dispose of unused insulin granules (66). However, these effects can only be extrapolated from other models, as studies on beta cells are lacking. Like other Enteroviruses, CVB3 hijacks the autophagy pathway to replicate into vesicles (67-69). Yet, the step at which CVB disturbs the autophagic processes remains controversial. Indeed, CVB has alternatively been reported to enhance the autophagic flux (68) or to drive autophagosome accumulation by inhibiting the fusion of autophagosomes with lysosomes and endosomes (67, 69). These discrepancies possibly reflect methodological differences, including cell models, infection conditions and autophagy readouts.

Summing up. Both mouse and human studies suggest that direct cytopathic effects of CVB infection are major contributors to beta-cell demise. Besides cytolysis, the strategies used by CVBs to hijack cellular pathways and favor its own replication impact beta-cell survival, identity and insulin secretion. The same is true for the pro-inflammatory cytokine release triggered by pattern recognition receptors, which enhances beta-cell apoptosis. Direct, immune-independent lytic effect on beta cells may secondarily trigger islet autoimmunity.

Indirect lytic effect of anti-viral immune responses on infected beta cells

Mouse studies. *Innate anti-viral responses.* CVB3 was unable to precipitate diabetes in NOD mice deficient for the NADPH oxidase, which display reduced superoxide production and

impaired M1 (pro-inflammatory) responses by macrophages (70). This underlines the importance of the inflammatory microenvironment driven by CVB3, independently of the direct, viral-mediated lysis. Conversely, diabetes protection may result from innate immune responses triggered by invariant natural killer T cells (71), which may limit viral spreading.

Antibody-mediated anti-viral responses. Mirroring data in the human (19), CVB3 infection induces neutralizing Abs in NOD mice (72). These Abs are transferred to the offspring of CVB3-infected NOD mice and protect the offspring from infection, and from diabetes development in *socs1*-transgenic NOD mice (72).

Effects of CVB infection on antigen processing and presentation. Beta-cell and immune-derived IFN responses increase surface MHC Class I expression and upregulate several genes of the antigen processing and presentation pathway. This results in an increased beta-cell visibility to islet-reactive, and possibly viral-reactive, CD8+ T cells. Indeed, CVB3 inoculation failed to exacerbate T-cell insulitis and diabetes onset in *tlr3*$^{-/-}$ NOD mice that mount impaired IFN responses (73), despite higher viral titers than in wild-type mice (74).

Following CVB4 infection, while necrosis occurs in neighboring acinar cells but not in islets, beta-cell engulfment by resident antigen-presenting cells (APCs) is observed in diabetes-resistant immunodeficient NOD/*scid* mice, and these APCs can prime diabetogenic islet-reactive BDC2.5 T cells *in vitro* (75). In addition, adoptive transfer of macrophages from CVB4-infected NOD/*scid* mice into NOD/BDC2.5 T-cell receptor (TCR)-transgenic animals triggers diabetes (75). This suggests that diabetes may result from the uptake of infected beta cells by APCs and subsequent presentation of islet antigens.

T-cell-mediated anti-viral responses. The mechanisms underlying diabetes protection in younger NOD mice following CVB infection are unclear, but increased TGF-β-producing regulatory T cells (Tregs) have been reported (76). In older NOD mice, data point to a role of CVB as a diabetes accelerator through bystander activation of islet-reactive effector T-cell responses and autoimmune beta-cell lysis. Only indirect clues are instead available on whether CVB-reactive T-cell responses are required to trigger diabetes. For instance, CVB4-infected NOD/*scid* mice develop high CVB4 titers but no diabetes (77), suggesting that CVB-mediated beta-cell lysis alone is not sufficient to trigger disease. Diabetes is instead triggered when CVB4-infected NOD/*scid* mice are adoptively transferred with islet-reactive BDC2.5 TCR-transgenic T cells (75), suggesting a requirement for islet-reactive but not CVB-reactive T cells.

Similarly, CVB4 induces diabetes in BDC2.5 TCR-transgenic NOD mice that are otherwise diabetes-resistant (78).

Collectively, the diabetes acceleration induced by CVB infection in NOD mice does not seem to exclusively rely on a direct, viral-mediated lytic effect on beta cells, but also on the presence of autoreactive T cells and the enhanced presentation of islet antigens by APCs through phagocytosis of infected beta cells. Anti-CVB T-cell responses seem instead dispensable and even protective.

Human studies. *Innate anti-viral immune responses (RM).* Apart from the beta-cell-autonomous pro-inflammatory response triggered by CVB infection through pattern recognition receptors, direct evidence about the innate immune responses that may be triggered by CVB infection in human T1D is very limited. This likely reflects experimental challenges: the question could be addressed either with in-vitro co-culture systems of human islets and relevant immune subsets or with histopathological studies, but neither have been attempted to date. Most histopathological studies have rather sought evidence of anti-viral responses in beta cells. A recent report from the DiViD study (79) documented that the pancreas from newly diagnosed living T1D donors harbored a small subset of VP1+ beta cells with markedly increased expression of the viral response protein kinase R. An increased islet expression of other viral response proteins, i.e. MDA5 and MxA, relative to non-diabetic controls was also noted.

Antibody-mediated anti-viral responses. CVB serology has often been used as an indirect readout of viral exposure (19). A recent work by the group of E. Bonifacio measured neutralizing anti-VP1 Abs to look at the magnitude of response according to islet aAb status (80). The results were striking: anti-VP1 Abs against all 6 major CVB serotypes were absent in children later developing early anti-insulin aAbs, which represent a fast-progressing T1D endotype (5, 81). This contrasted with the detection of anti-VP1 Abs in children without islet aAb seroconversion and in those positive for anti-glutamic acid decarboxylase (GAD) aAbs at a later age. A caveat of this study is that it did not provide direct evidence of prior CVB infection, hence the lack of anti-VP1 Abs could also reflect an absence of viral exposure. Nonetheless, these results may suggest that weak anti-CVB Ab responses predispose to islet autoimmunity. This scenario would help explain the epidemiological paradox that T1D incidence is higher where CVB exposure is lower. Indeed, CVB exposure drastically dropped over the last 40 years, while T1D incidence steadily increased (82). Also geographically (82, 83), the world highest T1D incidence of Finland contrasts with its very low CVB circulation.

This paradox has led to the 'Poliovirus hypothesis' (84, 85), postulating that this scenario may be similar to that of the beginning of the XX century, when reduced Poliovirus circulation due to improved sanitary conditions was paralleled by an increased incidence of its severe form, poliomyelitis (86). The explanation proposed is that a low frequency of CVB infection in the background population leads to decreased herd immunity and transmission of protective maternal Abs. Children are thus less protected during the first years of life and tend to develop higher-titer CVB infections, which favor viremia and viral spreading to vulnerable organs such as the heart or the pancreas (or motor neurons in the case of Poliovirus). It would be informative to know whether the incidence of CVB myocarditis has also increased during the last decades, but the improved diagnostic workup has identified CVB as a major etiologic agent (87) and introduces a major bias.

Effects of CVB infection on antigen processing and presentation. The CVB-induced disruption of membrane trafficking (decreased anterograde flux and increased endocytosis) that perturbs insulin secretion also represents an immune evasion mechanism. This disruption results in the internalization of surface MHC Class I molecules and limits presentation of newly-formed complexes, notably of viral peptides (64). It is plausible that internalized complexes are returned to the cell surface, circumventing natural killer cell detection (88). This mechanism likely involves the internalization of other surface molecules such as cytokine receptors.

Moreover, CVB infection distinctively alters the antigen processing and presentation pathway, as observed in CVB3-induced myocarditis models (89). Also in this case, self-limited viral myocarditis can lead to loss of tolerance to cardiac antigens and autoimmunity (37). Infection induces IFN-α expression, which upregulates MHC Class I and promotes the switch from the constitutive proteasome to immuno-proteasome (89). The immuno-proteasome has different cleavage preferences to process peptides for the MHC Class I pathway, thus generating more antigenic peptides for cell surface presentation. *In-vivo* studies revealed that CVB3-susceptible mouse strains display a longer and higher immuno-proteasome expression post-infection (89) along with higher expression of genes of the antigen processing and presentation pathway. However, most proteasomal enzymatic activities were reduced upon infection (89), suggesting the existence of other immune escape mechanisms to counteract the IFN-α effects. Similar processes might be at play in infected beta cells to limit anti-viral T-cell responses.

T-cell-mediated anti-viral responses. A major gap in knowledge concerns the features of anti-CVB T-cell responses, which is largely due to the lack of information about the viral epitopes recognized. A first report by M. Atkinson in 1994 (90) identified an antigenic CVB region by

first analyzing T-cell responses against GAD in unfractionated PBMCs stimulated with overlapping peptides. While these responses were detected in both T1D and healthy donors, a $GAD_{247-279}$ region was preferentially recognized in T1D and at-risk individuals compared to healthy controls. This region harbors a significant homology with the P2C protein of CVB (the so called PEVKEK region), and this homologous CVB sequence was recognized by the same donors. Although this suggests the possibility of epitope mimicry and T-cell cross-reactivity, formal proof at the single T-cell level and whether of the CD4+ or CD8+ subset was not provided. Indeed, a subsequent study by Schloot et al. (91) using CD4+ T-cell clones dismissed this possibility.

Anti-CVB CD4+ T-cell responses are likely to be elicited upon CVB infection, as they are required to drive B-cell activation and Ab production. Using individual recombinant CVB proteins, Varela-Calvino et al. (92) reported that the capsid proteins VP1, VP2 and VP3 were preferentially targeted by T cells (likely CD4+ T cells in the unfractionated PBMCs used in these assays), and that T1D patients harbored lower frequencies of proliferative responses against VP2 but stronger IFN-γ responses against VP3 and P2C. Another study (93) identified human CD4+ T-cell responses against Poliovirus VP2 and VP3 peptides selected from regions conserved across Enteroviruses.

CVB infection is also likely to recruit CD8+ T cells, as they are key players of viral clearance via the cytotoxic destruction of infected cells. Using HLA-A2 binding prediction algorithms, Varela-Calvino et al. (94) identified a $CVB4_{1137-1145}$ epitope (EVKEKHEFL) located in the same P2C region described by Atkinson et al. This peptide was naturally processed and presented by protein-pulsed APCs and recognized by IFN-γ-producing CD8+ T cells from 2 of 3 HLA-A2+ healthy donors tested. While T-cell lines raised against this epitope were cytotoxic against peptide-pulsed target B cells, there was no evidence of cross-reactivity with the homologous $GAD_{261-269}$ sequence (EVKEKGMAA). Using more sophisticated in-silico approaches, additional HLA-A2-restricted CVB epitopes were proposed, but cognate CD8+ T-cell responses detected by IFN-γ ELISpot were marginal (95).

Finally, like other viruses CVB can also mount mechanisms to escape T-cell recognition, such as downregulating surface HLA Class I expression (64). It is unknown whether HLA Class II expression is also downregulated. Another open question is whether CVB persistence may favor the induction of CD8+ T-cell exhaustion and, ultimately, poor anti-viral memory (96-98), and whether viral persistence is a common outcome following an acute CVB infection.

Summing up. There is very limited information about the anti-CVB CD4+ and CD8+ T-cell responses mounted upon infection that may secondarily cause beta-cell damage. This gap in knowledge reflects the lack of reliable epitopes to track them. Mouse studies support a role for autoimmune T cells, while anti-viral T cells may be protective rather than harmful.

Epitope mimicry

Mouse studies. CVB/GAD homologous peptides can be presented by the diabetes-predisposing I-A^{g7} MHC Class II allele of NOD mice (99). Although T-cell cross-reactivity was demonstrated by *in-vitro* functional assays, *in-vivo* evidence that CVB-reactive T cells can also recognize GAD peptides and amplify autoimmune anti-GAD T-cell responses or transfer diabetes is lacking. Indeed, T cells isolated from CVB4-infected NOD mice neither exhibited increased proliferation to GAD protein or homologous GAD/CVB peptide, nor triggered diabetes (78). Molecular similarities were also identified between the VP1 capsid protein of CVB and the T1D autoantigens tyrosine phosphatases IA-2 and IAR, and sera from NOD mice inoculated with CVB4 showed some cross-reactivity with an IAR peptide (100). However, whether such mechanism can trigger autoreactive T cells to attack beta cells has not been demonstrated.

Human studies. Despite some initial enthusiasm on the possibility of T-cell cross-reactivity between CVB and GAD peptides mapping to the PEVKEK region (90), subsequent reports dismissed this possibility for both CD4+ (91) and CD8+ T cells (94). Cross-reactive responses between a Rotavirus $VP7_{16-49}$ region and sequences from the IA-2 and GAD islet antigens encompassing both CD4+ and CD8+ epitopes have also been described (101).

Summing up. Although molecular mimicry is intriguing, the evidence for a CVB cross-reactivity with islet antigens that may underlie this hypothetical mechanism is limited and conflicting. Moreover, the evidence for a causal role of this cross-reactivity in triggering islet autoimmunity is missing altogether, in both humans and mice.

Implications for T1D prevention strategies

Despite the strength of evidence for an association between CVB infections and islet autoimmunity, demonstration of a cause-effect relationship is lacking, and is likely to remain elusive until tested in the human by removing this candidate environmental trigger (12). The

most effective way to achieve this would be through vaccination against CVB. Importantly, this could allow to prevent the most common manifestations of this infection (flu-like syndrome) and, more importantly, its rare but severe complications (myocarditis, encephalitis, meningitis). This is similar to what has been achieved by Rotavirus vaccination in several countries. Following this rationale, a formalin-inactivated non-adjuvanted CVB vaccine has been developed. Preclinical studies using a prototype CVB1 vaccine was tested in BALB/c and NOD mice (44). High titers of neutralizing Abs were induced. While non-immunized NOD mice displayed accelerated diabetes after CVB1 infection, this was not the case in vaccinated animals, suggesting that the vaccine itself does not accelerate diabetes development. A subsequent study in CVB1-infected *socs-1*-transgenic NOD mice, which are unable to respond to IFNs and thus develop diabetes due to massive beta-cell destruction, documented protection against both CVB1 infection and subsequent CVB1-induced diabetes (51). A hexavalent version of the vaccine comprising the 6 CVB serotypes was subsequently shown to induce strong neutralizing Ab responses without adjuvant in both mice and non-human primates, and provided immunity and protection against CVB-induced myocarditis and diabetes (102). In NOD mice, the vaccine did not accelerate spontaneous diabetes, while it delayed diabetes acceleration upon CVB1 infection (45).

On these grounds, a formalin-inactivated non-adjuvanted pentavalent intramuscular vaccine comprising the 5 most common serotypes (CVB1 to CVB5, i.e. barring the less prevalent CVB6) is undergoing a phase I safety trial in CVB-seronegative healthy adults (NCT04690426). Early timing of vaccination will be critical for subsequent trials in order to intervene before CVB exposure, e.g. starting at 2 months of age as for the inactivated polio (Salk) vaccine. This will also require an excellent safety profile. In this perspective, it is important to gain further insights into the mechanisms by which CVB infection may trigger beta-cell autoimmunity. Such mechanisms hold implications for the clinical benefit that may be expected (**Fig. 1**). If beta-cell destruction is mainly provoked by the secondary lytic effect of anti-viral immune responses on infected beta cells, vaccination may increase such responses and accelerate beta-cell destruction. The same could be true if epitope mimicry mechanisms contribute to this destruction, unless the viral sequences at play are excised from the vaccine constructs. While this possibility assumes that CVB may still reach the pancreas in vaccinated individuals, preclinical studies demonstrate that the vaccine efficiently prevents infection ab initio, including viremia and systemic spreading (44, 45, 102). Conversely, the scenario of a primary lytic effect of CVB on infected beta cells leading to self-antigen release and

autoimmune priming, possibly associated with poor anti-CVB immune responses, would lend a strong rationale for boosting these responses by vaccination.

Conclusions and Perspectives

Recent advances in our understanding of the association between CVB infections and islet autoimmunity, and between Epstein-Barr virus and multiple sclerosis, invite us to move one step further to definitely prove or disprove causality and explore preventative interventions through anti-viral vaccination. In parallel, a better knowledge of the immune responses elicited by CVB infections and vaccines will allow us to maximize the risk/benefit ratio and to optimize vaccination strategies by providing immune monitoring tools and surrogate markers to predict protection.

Acknowledgements. Figures 1 was created with elements from Servier Medical Art (smart.servier.com).

Funding. Work performed in the Laboratory of R.M. and M.F.T. was supported by JDRF nPOD-Virus grant 3-SRA-2017-492-A-N. The laboratory of R.M. and S.Y. is supported by The Leona M. and Harry B. Helmsley Charitable Trust (Helmsley #1901-03689), the *Agence Nationale de la Recherche* (ANR-19-CE15-0014-01) and the *Fondation pour la Recherche Medicale* (EQU20193007831); and by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, the European Federation of Pharmaceutical Industries and Associations, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust. The laboratory of M.F.T. is supported by The Swedish Child Diabetes Foundation, The Swedish Diabetes Foundation, The Swedish Research Council (#2020-02969), The Swedish Heart and Lung Foundation (#20200682), Novo Nordic Foundation, Denmark, and Karolinska Institutet including the Strategic Research Program in Diabetes (MFT).

Duality of Interest. R.M. receives research funding from Provention Bio Inc. M.F.–T. serves on the scientific advisory board of Provention Bio Inc. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. A.C., F.V., S.Y. and R.M. wrote manuscript. M.F.T. contributed to discussion and reviewed/edited manuscript. R.M. is the guarantor of this work.

References

1. Craig ME, Kim KW, Isaacs SR, Penno MA, Hamilton-Williams EE, Couper JJ, Rawlinson WD: Early-life factors contributing to type 1 diabetes. Diabetologia 2019;62:1823-1834
2. Gillespie KM, Bain SC, Barnett AH, Bingley PJ, Christie MR, Gill GV, Gale EA: The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. Lancet 2004;364:1699-1700
3. Ziegler AG, Bonifacio E, Group B-BS: Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes. Diabetologia 2012;55:1937-1943
4. Parikka V, Nanto-Salonen K, Saarinen M, Simell T, Ilonen J, Hyoty H, Veijola R, Knip M, Simell O: Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk. Diabetologia 2012;55:1926-1936
5. Carré A, Richardson SJ, Larger E, Mallone R: Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? Diabetologia 2021;64:15-25
6. Hyöty H: Viruses in type 1 diabetes. Pediatr Diabetes 2016;17 Suppl 22:56-64
7. Nekoua MP, Alidjinou EK, Hober D: Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus. Nat Rev Endocrinol 2022;
8. Bjornevik K, Cortese M, Healy BC, Kuhle J, Mina MJ, Leng Y, Elledge SJ, Niebuhr DW, Scher AI, Munger KL, Ascherio A: Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. Science 2022;375:296-301
9. Horwitz RI, Hayes-Conroy A, Singer BH, Cullen MR, Badal K, Sim I: Falling down the biological rabbit hole: Epstein-Barr virus, biography, and multiple sclerosis. The Journal of Clinical Investigation 2022;132
10. Lanz TV, Brewer RC, Ho PP, Moon JS, Jude KM, Fernandez D, Fernandes RA, Gomez AM, Nadj GS, Bartley CM, Schubert RD, Hawes IA, Vazquez SE, Iyer M, Zuchero JB, Teegen B, Dunn JE, Lock CB, Kipp LB, Cotham VC, Ueberheide BM, Aftab BT, Anderson MS, DeRisi JL, Wilson MR, Bashford-Rogers RJM, Platten M, Garcia KC, Steinman L, Robinson WH: Clonally Expanded B Cells in Multiple Sclerosis Bind EBV EBNA1 and GlialCAM. Nature 2022;
11. Kuchroo VK, Weiner HL: How does Epstein-Barr virus trigger MS? Immunity 2022;55:390-392
12. Dunne JL, Richardson SJ, Atkinson MA, Craig ME, Dahl-Jorgensen K, Flodstrom-Tullberg M, Hyoty H, Insel RA, Lernmark A, Lloyd RE, Morgan NG, Pugliese A: Rationale for enteroviral vaccination and antiviral therapies in human type 1 diabetes. Diabetologia 2019;62:744-753
13. Richardson SJ, Morgan NG: Enteroviral infections in the pathogenesis of type 1 diabetes: new insights for therapeutic intervention. Curr Opin Pharmacol 2018;43:11-19
14. Rodriguez-Calvo T, Sabouri S, Anquetil F, von Herrath MG: The viral paradigm in type 1 diabetes: Who are the main suspects? Autoimmunity reviews 2016;15:964-969
15. Chapman NM, Kim KS: Persistent coxsackievirus infection: enterovirus persistence in chronic myocarditis and dilated cardiomyopathy. Curr Top Microbiol Immunol 2008;323:275-292
16. Feuer R, Ruller CM, An N, Tabor-Godwin JM, Rhoades RE, Maciejewski S, Pagarigan RR, Cornell CT, Crocker SJ, Kiosses WB, Pham-Mitchell N, Campbell IL, Whitton JL: Viral persistence and chronic immunopathology in the adult central nervous system following Coxsackievirus infection during the neonatal period. J Virol 2009;83:9356-9369
17. Marjomaki V, Flodstrom-Tullberg M: Coxsackie B virus. Trends Microbiol 2022;30:606-607

18. Yeung WC, Rawlinson WD, Craig ME: Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies. BMJ 2011;342:d35
19. Laitinen OH, Honkanen H, Pakkanen O, Oikarinen S, Hankaniemi MM, Huhtala H, Ruokoranta T, Lecouturier V, Andre P, Harju R, Virtanen SM, Lehtonen J, Almond JW, Simell T, Simell O, Ilonen J, Veijola R, Knip M, Hyoty H: Coxsackievirus B1 Is Associated With Induction of beta-Cell Autoimmunity That Portends Type 1 Diabetes. Diabetes 2014;63:446-455
20. Vehik K, Lynch KF, Wong MC, Tian X, Ross MC, Gibbs RA, Ajami NJ, Petrosino JF, Rewers M, Toppari J, Ziegler AG, She JX, Lernmark A, Akolkar B, Hagopian WA, Schatz DA, Krischer JP, Hyoty H, Lloyd RE, Group TS: Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. Nat Med 2019;25:1865-1872
21. Cabrera-Rode E, Sarmiento L, Tiberti C, Molina G, Barrios J, Hernández D, Díaz-Horta O, Di Mario U: Type 1 diabetes islet associated antibodies in subjects infected by echovirus 16. Diabetologia 2003;46:1348-1353
22. Cabrera-Rode E, Sarmiento L, Molina G, Pérez C, Arranz C, Galvan JA, Prieto M, Barrios J, Palomera R, Fonseca M, Mas P, Díaz-Díaz O, Díaz-Horta O: Islet cell related antibodies and type 1 diabetes associated with echovirus 30 epidemic: a case report. Journal of medical virology 2005;76:373-377
23. Sarmiento L, Galvan JA, Cabrera-Rode E, Aira L, Correa C, Sariego S, Fonseca M, Cubas-Dueñas I, Hung LH, Resik S, Cilio CM: Type 1 diabetes associated and tissue transglutaminase autoantibodies in patients without type 1 diabetes and coeliac disease with confirmed viral infections. Journal of medical virology 2012;84:1049-1053
24. Dotta F, Censini S, van Halteren AG, Marselli L, Masini M, Dionisi S, Mosca F, Boggi U, Muda AO, Prato SD, Elliott JF, Covacci A, Rappuoli R, Roep BO, Marchetti P: Coxsackie B4 virus infection of beta cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. Proc Natl Acad Sci USA 2007;104:5115-5120
25. Richardson SJ, Willcox A, Bone AJ, Foulis AK, Morgan NG: The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. Diabetologia 2009;52:1143-1151
26. Richardson SJ, Leete P, Bone AJ, Foulis AK, Morgan NG: Expression of the enteroviral capsid protein VP1 in the islet cells of patients with type 1 diabetes is associated with induction of protein kinase R and downregulation of Mcl-1. Diabetologia 2013;56:185-193
27. Krogvold L, Edwin B, Buanes T, Frisk G, Skog O, Anagandula M, Korsgren O, Undlien D, Eike MC, Richardson SJ, Leete P, Morgan NG, Oikarinen S, Oikarinen M, Laiho JE, Hyoty H, Ludvigsson J, Hanssen KF, Dahl-Jorgensen K: Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes. Diabetes 2015;64:1682-1687
28. Busse N, Paroni F, Richardson SJ, Laiho JE, Oikarinen M, Frisk G, Hyoty H, de Koning E, Morgan NG, Maedler K: Detection and localization of viral infection in the pancreas of patients with type 1 diabetes using short fluorescently-labelled oligonucleotide probes. Oncotarget 2017;8:12620-12636
29. Netanyah E, Calafatti M, Arvastsson J, Cabrera-Rode E, Cilio CM, Sarmiento L: Extracellular Vesicles Released by Enterovirus-Infected EndoC-βH1 Cells Mediate Non-Lytic Viral Spread. Microorganisms 2020;8
30. Zanone MM, Favaro E, Conaldi PG, Greening J, Bottelli A, Perin PC, Klein NJ, Peakman M, Camussi G: Persistent infection of human microvascular endothelial cells by coxsackie B viruses induces increased expression of adhesion molecules. JImmunol 2003;171:438-446
31. Kim KS, Tracy S, Tapprich W, Bailey J, Lee CK, Kim K, Barry WH, Chapman NM: 5'-Terminal deletions occur in coxsackievirus B3 during replication in murine hearts and cardiac

myocyte cultures and correlate with encapsidation of negative-strand viral RNA. J Virol 2005;79:7024-7041
32. Kallionpaa H, Elo LL, Laajala E, Mykkanen J, Ricano-Ponce I, Vaarma M, Laajala TD, Hyoty H, Ilonen J, Veijola R, Simell T, Wijmenga C, Knip M, Lahdesmaki H, Simell O, Lahesmaa R: Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility. Diabetes 2014;63:2402-2414
33. Ferreira RC, Guo H, Coulson RM, Smyth DJ, Pekalski ML, Burren OS, Cutler AJ, Doecke JD, Flint S, McKinney EF, Lyons PA, Smith KG, Achenbach P, Beyerlein A, Dunger DB, Clayton DG, Wicker LS, Todd JA, Bonifacio E, Wallace C, Ziegler AG: A type I interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. Diabetes 2014;63:2538-2550
34. Chehadeh W, Kerr-Conte J, Pattou F, Alm G, Lefebvre J, Wattre P, Hober D: Persistent infection of human pancreatic islets by coxsackievirus B is associated with alpha interferon synthesis in beta cells. J Virol 2000;74:10153-10164
35. See DM, Tilles JG: Pathogenesis of virus-induced diabetes in mice. J Infect Dis 1995;171:1131-1138
36. Horwitz MS, Ilic A, Fine C, Rodriguez E, Sarvetnick N: Coxsackievirus-mediated hyperglycemia is enhanced by reinfection and this occurs independent of T cells. Virology 2003;314:510-520
37. Rose NR: Myocarditis: infection versus autoimmunity. J Clin Immunol 2009;29:730-737
38. Toniolo A, Onodera T, Yoon JW, Notkins AL: Induction of diabetes by cumulative environmental insults from viruses and chemicals. Nature 1980;288:383-385
39. Benkahla MA, Sane F, Bertin A, Vreulx AC, Elmastour F, Jaidane H, Desailloud R, Hober D: Impact of coxsackievirus-B4E2 combined with a single low dose of streptozotocin on pancreas of outbred mice: investigation of viral load, pathology and inflammation. Scientific reports 2019;9:10080
40. Gallagher GR, Brehm MA, Finberg RW, Barton BA, Shultz LD, Greiner DL, Bortell R, Wang JP: Viral infection of engrafted human islets leads to diabetes. Diabetes 2015;64:1358-1369
41. Bernard H, Teijeiro A, Chaves-Perez A, Perna C, Satish B, Novials A, Wang JP, Djouder N: Coxsackievirus B Type 4 Infection in beta Cells Downregulates the Chaperone Prefoldin URI to Induce a MODY4-like Diabetes via Pdx1 Silencing. Cell Rep Med 2020;1:100125
42. Serreze DV, Ottendorfer EW, Ellis TM, Gauntt CJ, Atkinson MA: Acceleration of type 1 diabetes by a coxsackievirus infection requires a preexisting critical mass of autoreactive T-cells in pancreatic islets. Diabetes 2000;49:708-711
43. Drescher KM, Kono K, Bopegamage S, Carson SD, Tracy S: Coxsackievirus B3 infection and type 1 diabetes development in NOD mice: insulitis determines susceptibility of pancreatic islets to virus infection. Virology 2004;329:381-394
44. Larsson PG, Lakshmikanth T, Laitinen OH, Utorova R, Jacobson S, Oikarinen M, Domsgen E, Koivunen MR, Chaux P, Devard N, Lecouturier V, Almond J, Knip M, Hyöty H, Flodström-Tullberg M: A preclinical study on the efficacy and safety of a new vaccine against Coxsackievirus B1 reveals no risk for accelerated diabetes development in mouse models. Diabetologia 2015;58:346-354
45. Stone VM, Butrym M, Hankaniemi MM, Sioofy-Khojine A-B, Hytönen VP, Hyöty H, Flodström-Tullberg M: Coxsackievirus B Vaccines Prevent Infection-Accelerated Diabetes in NOD Mice and Have No Disease-Inducing Effect. Diabetes 2021;70:2871-2878
46. Kanno T, Kim K, Kono K, Drescher KM, Chapman NM, Tracy S: Group B coxsackievirus diabetogenic phenotype correlates with replication efficiency. J Virol 2006;80:5637-5643
47. Tracy S, Drescher KM, Chapman NM, Kim KS, Carson SD, Pirruccello S, Lane PH, Romero JR, Leser JS: Toward testing the hypothesis that group B coxsackieviruses (CVB)

trigger insulin-dependent diabetes: inoculating nonobese diabetic mice with CVB markedly lowers diabetes incidence. J Virol 2002;76:12097-12111
48. Lalwani A, Warren J, Liuwantara D, Hawthorne WJ, O'Connell PJ, Gonzalez FJ, Stokes RA, Chen J, Laybutt DR, Craig ME, Swarbrick MM, King C, Gunton JE: beta Cell Hypoxia-Inducible Factor-1alpha Is Required for the Prevention of Type 1 Diabetes. Cell reports 2019;27:2370-2384 e2376
49. Flodstrom M, Maday A, Balakrishna D, Cleary MM, Yoshimura A, Sarvetnick N: Target cell defense prevents the development of diabetes after viral infection. Nat Immunol 2002;3:373-382
50. Flodstrom M, Tsai D, Fine C, Maday A, Sarvetnick N: Diabetogenic potential of human pathogens uncovered in experimentally permissive beta-cells. Diabetes 2003;52:2025-2034
51. Stone VM, Hankaniemi MM, Svedin E, Sioofy-Khojine A, Oikarinen S, Hyoty H, Laitinen OH, Hytonen VP, Flodstrom-Tullberg M: A Coxsackievirus B vaccine protects against virus-induced diabetes in an experimental mouse model of type 1 diabetes. Diabetologia 2018;61:476-481
52. Roivainen M, Rasilainen S, Ylipaasto P, Nissinen R, Ustinov J, Bouwens L, Eizirik DL, Hovi T, Otonkoski T: Mechanisms of coxsackievirus-induced damage to human pancreatic beta-cells. J Clin Endocrinol Metab 2000;85:432-440
53. Feuer R, Mena I, Pagarigan R, Slifka MK, Whitton JL: Cell cycle status affects coxsackievirus replication, persistence, and reactivation in vitro. J Virol 2002;76:4430-4440
54. Fan Y, Sanyal S, Bruzzone R: Breaking Bad: How Viruses Subvert the Cell Cycle. Front Cell Infect Microbiol 2018;8:396
55. Wang ZY, Zhong T, Wang Y, Song FM, Yu XF, Xing LP, Zhang WY, Yu JH, Hua SC, Yu XF: Human Enterovirus 68 Interferes with the Host Cell Cycle to Facilitate Viral Production. Front Cell Infect Microbiol 2017;7:29
56. Ylipaasto P, Smura T, Gopalacharyulu P, Paananen A, Seppanen-Laakso T, Kaijalainen S, Ahlfors H, Korsgren O, Lakey JR, Lahesmaa R, Piemonti L, Oresic M, Galama J, Roivainen M: Enterovirus-induced gene expression profile is critical for human pancreatic islet destruction. Diabetologia 2012;55:3273-3283
57. Oshima M, Knoch KP, Diedisheim M, Petzold A, Cattan P, Bugliani M, Marchetti P, Choudhary P, Huang GC, Bornstein SR, Solimena M, Albagli-Curiel O, Scharfmann R: Virus-like infection induces human beta cell dedifferentiation. JCI insight 2018;3
58. Hodik M, Skog O, Lukinius A, Isaza-Correa JM, Kuipers J, Giepmans BN, Frisk G: Enterovirus infection of human islets of Langerhans affects beta-cell function resulting in disintegrated islets, decreased glucose stimulated insulin secretion and loss of Golgi structure. BMJ Open Diabetes Res Care 2016;4:e000179
59. van Kuppeveld FJ, Hoenderop JG, Smeets RL, Willems PH, Dijkman HB, Galama JM, Melchers WJ: Coxsackievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. Embo J 1997;16:3519-3532
60. Fung TS, Torres J, Liu DX: The Emerging Roles of Viroporins in ER Stress Response and Autophagy Induction during Virus Infection. Viruses 2015;7:2834-2857
61. Laitinen OH, Svedin E, Kapell S, Nurminen A, Hytonen VP, Flodstrom-Tullberg M: Enteroviral proteases: structure, host interactions and pathogenicity. Reviews in medical virology 2016;26:251-267
62. Mallone R, Eizirik DL: Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? Diabetologia 2020;63:1999-2006
63. Colli ML, Paula FM, Marselli L, Marchetti P, Roivainen M, Eizirik DL, Op de Beeck A: Coxsackievirus B Tailors the Unfolded Protein Response to Favour Viral Amplification in Pancreatic beta Cells. Journal of innate immunity 2019;11:375-390

64. Cornell CT, Kiosses WB, Harkins S, Whitton JL: Coxsackievirus B3 proteins directionally complement each other to downregulate surface major histocompatibility complex class I. J Virol 2007;81:6785-6797
65. Knoch KP, Nath-Sain S, Petzold A, Schneider H, Beck M, Wegbrod C, Sonmez A, Munster C, Friedrich A, Roivainen M, Solimena M: PTBP1 is required for glucose-stimulated cap-independent translation of insulin granule proteins and Coxsackieviruses in beta cells. Molecular metabolism 2014;3:518-530
66. Carre A, Mallone R: Making Insulin and Staying Out of Autoimmune Trouble: The Beta-Cell Conundrum. Front Immunol 2021;12:639682
67. Kemball CC, Alirezaei M, Flynn CT, Wood MR, Harkins S, Kiosses WB, Whitton JL: Coxsackievirus infection induces autophagy-like vesicles and megaphagosomes in pancreatic acinar cells in vivo. J Virol 2010;84:12110-12124
68. Shi X, Chen Z, Tang S, Wu F, Xiong S, Dong C: Coxsackievirus B3 infection induces autophagic flux, and autophagosomes are critical for efficient viral replication. Arch Virol 2016;161:2197-2205
69. Mohamud Y, Shi J, Qu J, Poon T, Xue YC, Deng H, Zhang J, Luo H: Enteroviral Infection Inhibits Autophagic Flux via Disruption of the SNARE Complex to Enhance Viral Replication. Cell reports 2018;22:3292-3303
70. Burg AR, Das S, Padgett LE, Koenig ZE, Tse HM: Superoxide Production by NADPH Oxidase Intensifies Macrophage Antiviral Responses during Diabetogenic Coxsackievirus Infection. J Immunol 2018;200:61-70
71. Ghazarian L, Diana J, Beaudoin L, Larsson PG, Puri RK, van Rooijen N, Flodstrom-Tullberg M, Lehuen A: Protection against type 1 diabetes upon Coxsackievirus B4 infection and iNKT-cell stimulation: role of suppressive macrophages. Diabetes 2013;62:3785-3796
72. Larsson PG, Lakshmikanth T, Svedin E, King C, Flodstrom-Tullberg M: Previous maternal infection protects offspring from enterovirus infection and prevents experimental diabetes development in mice. Diabetologia 2013;56:867-874
73. Seleme MC, Lei W, Burg AR, Goh KY, Metz A, Steele C, Tse HM: Dysregulated TLR3-dependent signaling and innate immune activation in superoxide-deficient macrophages from nonobese diabetic mice. Free Radic Biol Med 2012;52:2047-2056
74. McCall KD, Thuma JR, Courreges MC, Benencia F, James CB, Malgor R, Kantake N, Mudd W, Denlinger N, Nolan B, Wen L, Schwartz FL: Toll-like receptor 3 is critical for coxsackievirus B4-induced type 1 diabetes in female NOD mice. Endocrinology 2015;156:453-461
75. Horwitz MS, Ilic A, Fine C, Balasa B, Sarvetnick N: Coxsackieviral-mediated diabetes: induction requires antigen-presenting cells and is accompanied by phagocytosis of beta cells. Clin Immunol 2004;110:134-144
76. Filippi CM, Estes EA, Oldham JE, von Herrath MG: Immunoregulatory mechanisms triggered by viral infections protect from type 1 diabetes in mice. JClinInvest 2009;
77. Horwitz MS, Krahl T, Fine C, Lee J, Sarvetnick N: Protection from lethal coxsackievirus-induced pancreatitis by expression of gamma interferon. J Virol 1999;73:1756-1766
78. Horwitz MS, Bradley LM, Harbertson J, Krahl T, Lee J, Sarvetnick N: Diabetes induced by Coxsackie virus: initiation by bystander damage and not molecular mimicry. Nat Med 1998;4:781-785
79. Krogvold L, Leete P, Mynarek IM, Russell MA, Gerling IC, Lenchik NI, Mathews C, Richardson SJ, Morgan NG, Dahl-Jørgensen K: Detection of Antiviral Tissue Responses and Increased Cell Stress in the Pancreatic Islets of Newly Diagnosed Type 1 Diabetes Patients: Results From the DiViD Study. Frontiers in endocrinology 2022;13:881997

80. Ashton MP, Eugster A, Walther D, Daehling N, Riethausen S, Kuehn D, Klingel K, Beyerlein A, Zillmer S, Ziegler AG, Bonifacio E: Incomplete immune response to coxsackie B viruses associates with early autoimmunity against insulin. Scientific reports 2016;6:32899
81. Deligne C, You S, Mallone R: Personalized Immunotherapies for Type 1 Diabetes: Who, What, When, and How? J Pers Med 2022;12
82. Viskari H, Ludvigsson J, Uibo R, Salur L, Marciulionyte D, Hermann R, Soltesz G, Füchtenbusch M, Ziegler AG, Kondrashova A, Romanov A, Kaplan B, Laron Z, Koskela P, Vesikari T, Huhtala H, Knip M, Hyöty H: Relationship between the incidence of type 1 diabetes and maternal enterovirus antibodies: time trends and geographical variation. Diabetologia 2005;48:1280-1287
83. Viskari H, Ludvigsson J, Uibo R, Salur L, Marciulionyte D, Hermann R, Soltesz G, Füchtenbusch M, Ziegler AG, Kondrashova A, Romanov A, Knip M, Hyöty H: Relationship between the incidence of type 1 diabetes and enterovirus infections in different European populations: results from the EPIVIR project. Journal of medical virology 2004;72:610-617
84. Viskari HR, Koskela P, Lonnrot M, Luonuansuu S, Reunanen A, Baer M, Hyoty H: Can enterovirus infections explain the increasing incidence of type 1 diabetes? Diabetes Care 2000;23:414-416
85. Zinkernagel RM: Maternal antibodies, childhood infections, and autoimmune diseases. N Engl J Med 2001;345:1331-1335
86. Nathanson N, Kew OM: From emergence to eradication: the epidemiology of poliomyelitis deconstructed. American journal of epidemiology 2010;172:1213-1229
87. Pollack A, Kontorovich AR, Fuster V, Dec GW: Viral myocarditis--diagnosis, treatment options, and current controversies. Nat Rev Cardiol 2015;12:670-680
88. Huhn MH, Hultcrantz M, Lind K, Ljunggren HG, Malmberg KJ, Flodstrom-Tullberg M: IFN-gamma production dominates the early human natural killer cell response to Coxsackievirus infection. Cellular microbiology 2008;10:426-436
89. Szalay G, Meiners S, Voigt A, Lauber J, Spieth C, Speer N, Sauter M, Kuckelkorn U, Zell A, Klingel K, Stangl K, Kandolf R: Ongoing coxsackievirus myocarditis is associated with increased formation and activity of myocardial immunoproteasomes. Am J Pathol 2006;168:1542-1552
90. Atkinson MA, Bowman MA, Campbell L, Darrow BL, Kaufman DL, Maclaren NK: Cellular immunity to a determinant common to glutamate decarboxylase and coxsackie virus in insulin-dependent diabetes. J Clin Invest 1994;94:2125-2129
91. Schloot NC, Willemen SJ, Duinkerken G, Drijfhout JW, de Vries RR, Roep BO: Molecular mimicry in type 1 diabetes mellitus revisited: T-cell clones to GAD65 peptides with sequence homology to Coxsackie or proinsulin peptides do not crossreact with homologous counterpart. Hum Immunol 2001;62:299-309
92. Varela-Calvino R, Ellis R, Sgarbi G, Dayan CM, Peakman M: Characterization of the T-cell response to coxsackievirus B4: evidence that effector memory cells predominate in patients with type 1 diabetes. Diabetes 2002;51:1745-1753
93. Cello J, Strannegård O, Svennerholm B: A study of the cellular immune response to enteroviruses in humans: identification of cross-reactive T cell epitopes on the structural proteins of enteroviruses. J Gen Virol 1996;77 (Pt 9):2097-2108
94. Varela-Calvino R, Skowera A, Arif S, Peakman M: Identification of a naturally processed cytotoxic CD8 T-cell epitope of coxsackievirus B4, presented by HLA-A2.1 and located in the PEVKEK region of the P2C nonstructural protein. J Virol 2004;78:13399-13408
95. Kundu R, Knight R, Dunga M, Peakman M: In silico and ex vivo approaches indicate immune pressure on capsid and non-capsid regions of coxsackie B viruses in the human system. PLoS One 2018;13:e0199323

96. McLane LM, Abdel-Hakeem MS, Wherry EJ: CD8 T Cell Exhaustion During Chronic Viral Infection and Cancer. Annual Review of Immunology 2019;37:457-495
97. Blank CU, Haining WN, Held W, Hogan PG, Kallies A, Lugli E, Lynn RC, Philip M, Rao A, Restifo NP, Schietinger A, Schumacher TN, Schwartzberg PL, Sharpe AH, Speiser DE, Wherry EJ, Youngblood BA, Zehn D: Defining 'T cell exhaustion'. Nat Rev Immunol 2019;19:665-674
98. Zehn D, Thimme R, Lugli E, de Almeida GP, Oxenius A: 'Stem-like' precursors are the fount to sustain persistent CD8(+) T cell responses. Nat Immunol 2022;23:836-847
99. Tian J, Lehmann PV, Kaufman DL: T cell cross-reactivity between coxsackievirus and glutamate decarboxylase is associated with a murine diabetes susceptibility allele. J Exp Med 1994;180:1979-1984
100. Harkonen T, Lankinen H, Davydova B, Hovi T, Roivainen M: Enterovirus infection can induce immune responses that cross-react with beta-cell autoantigen tyrosine phosphatase IA-2/IAR. Journal of medical virology 2002;66:340-350
101. Honeyman MC, Stone NL, Falk BA, Nepom G, Harrison LC: Evidence for molecular mimicry between human T cell epitopes in rotavirus and pancreatic islet autoantigens. J Immunol 2010;184:2204-2210
102. Stone VM, Hankaniemi MM, Laitinen OH, Sioofy-Khojine AB, Lin A, Diaz Lozano IM, Mazur MA, Marjomaki V, Lore K, Hyoty H, Hytonen VP, Flodstrom-Tullberg M: A hexavalent Coxsackievirus B vaccine is highly immunogenic and has a strong protective capacity in mice and nonhuman primates. Science advances 2020;6:eaaz2433

Diabetologia (2021) 64:15–25
https://doi.org/10.1007/s00125-020-05298-y

REVIEW

Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset?

Alexia Carré[1] · Sarah J. Richardson[2] · Etienne Larger[1,3] · Roberto Mallone[1,3]

Received: 6 June 2020 / Accepted: 20 August 2020 / Published online: 21 October 2020
© The Author(s) 2020

Abstract

Available evidence provides arguments both for and against a primary pathogenic role for T cells in human type 1 diabetes. Genetic susceptibility linked to HLA Class II lends strong support. Histopathology documents HLA Class I hyperexpression and islet infiltrates dominated by $CD8^+$ T cells. While both hallmarks are near absent in autoantibody-positive donors, the variable insulitis and residual beta cells of recent-onset donors suggests the existence of a younger-onset endotype with more aggressive autoimmunity and an older-onset endotype with more vulnerable beta cells. Functional arguments from ex vivo and in vitro human studies and in vivo 'humanised' mouse models are instead neutral or against a T cell role. Clinical support is provided by the appearance of islet autoantibodies before disease onset. The faster C-peptide loss and superior benefits of immunotherapies in individuals with younger-onset type 1 diabetes reinforce the view of age-related endotypes. Clarifying the relative role of T cells will require technical advances in the identification of their target antigens, in their detection and phenotyping in the blood and pancreas, and in the study of the T cell/beta cell crosstalk. Critical steps toward this goal include the understanding of the link with environmental triggers, the description of T cell changes along the natural history of disease, and their relationship with age and the 'benign' islet autoimmunity of healthy individuals.

Keywords Autoantibodies · Autoimmunity · Beta cells · Endotypes · HLA · Immunotherapy · Insulitis · Islets · Pancreas · Review · T cells

Abbreviations

DiViD	Diabetes Virus Detection Study
EADB	Exeter Archival Diabetes Biobank
FOXP3	Forkhead box P3
HLA-I	HLA Class I
HLA-II	HLA Class II
ICI	Insulin-containing islet
IGRP	Islet-specific glucose-6-phosphatase catalytic-subunit related protein
MHC-I	MHC Class I
nPOD	Network for Pancreatic Organ Donors with Diabetes
TCR	T cell receptor
T1DE1	Type 1 diabetes endotype 1
T1DE2	Type 1 diabetes endotype 2
Treg	Regulatory T cell

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00125-020-05298-y) contains a slideset of the figures for download, which is available to authorised users.

✉ Roberto Mallone
roberto.mallone@inserm.fr

✉ Sarah J. Richardson
s.richardson@exeter.ac.uk

[1] Université de Paris, Institut Cochin, CNRS, INSERM, Paris, France

[2] Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, UK

[3] Assistance Publique Hôpitaux de Paris, Hôpitaux Universitaires de Paris Centre-Université de Paris, Cochin Hospital, Service de Diabétologie et Immunologie Clinique, Paris, France

Introduction

While type 1 diabetes is described as a T cell-mediated autoimmune disease, a more holistic view comprising the dialogue between T cell aggressors and beta cell targets is gaining credit. On one hand, this novel view underlines the active pathogenic role played by beta cells (reviewed in [1]) and on the other, it calls for a critical reappraisal of the role of T cells, which we address here. We contend that evidence for a primary or exclusive role of T cells is variable and probably reflects a balance with mechanisms of beta cell dysfunction that

underlie different disease subtypes (endotypes) [2]. These endotypes may thus feature T cells as either 'lead culprits' or 'partners in crime' for beta cell destruction. We will review arguments for and against T cell-mediated mechanisms (summarised in Fig. 1), link them to age-related endotypes, and propose future directions to settle these questions.

Genetic evidence

For Type 1 diabetes susceptibility and protection are strongly associated with HLA Class II (HLA-II) and, to lesser extent, HLA Class I (HLA-I) loci [3], with *HLA-DQB1*06:02* exerting a dominant protection. Given the antigen-presenting function of HLA molecules, these associations support a role for T cells. Several other disease-associated gene variants regulate T cell responses (e.g. *PTPN22*, *IL2RA*, *CTLA4*) [3]. Moreover, disease-associated, and likely causal, genetic variants are enriched in open chromatin (i.e. accessible for transcription) specifically in immune cells, particularly in $CD4^+$ effector T cells [4]. The rare genetic poly-autoimmune syndromes immune dysregulation, polyendocrinopathy, enteropathy, X-linked (IPEX) syndrome (*FOXP3* mutations) and autoimmune polyendocrinopathy, candidiasis, ectodermal dystrophy (APECED) syndrome (*AIRE* mutations) affect T cell tolerance mechanisms and can comprise type 1 diabetes. Genetically imprinted, age-related disease endotypes may exist, as younger (<13 months old) children more frequently carry *HLA-DR4/DQ8*

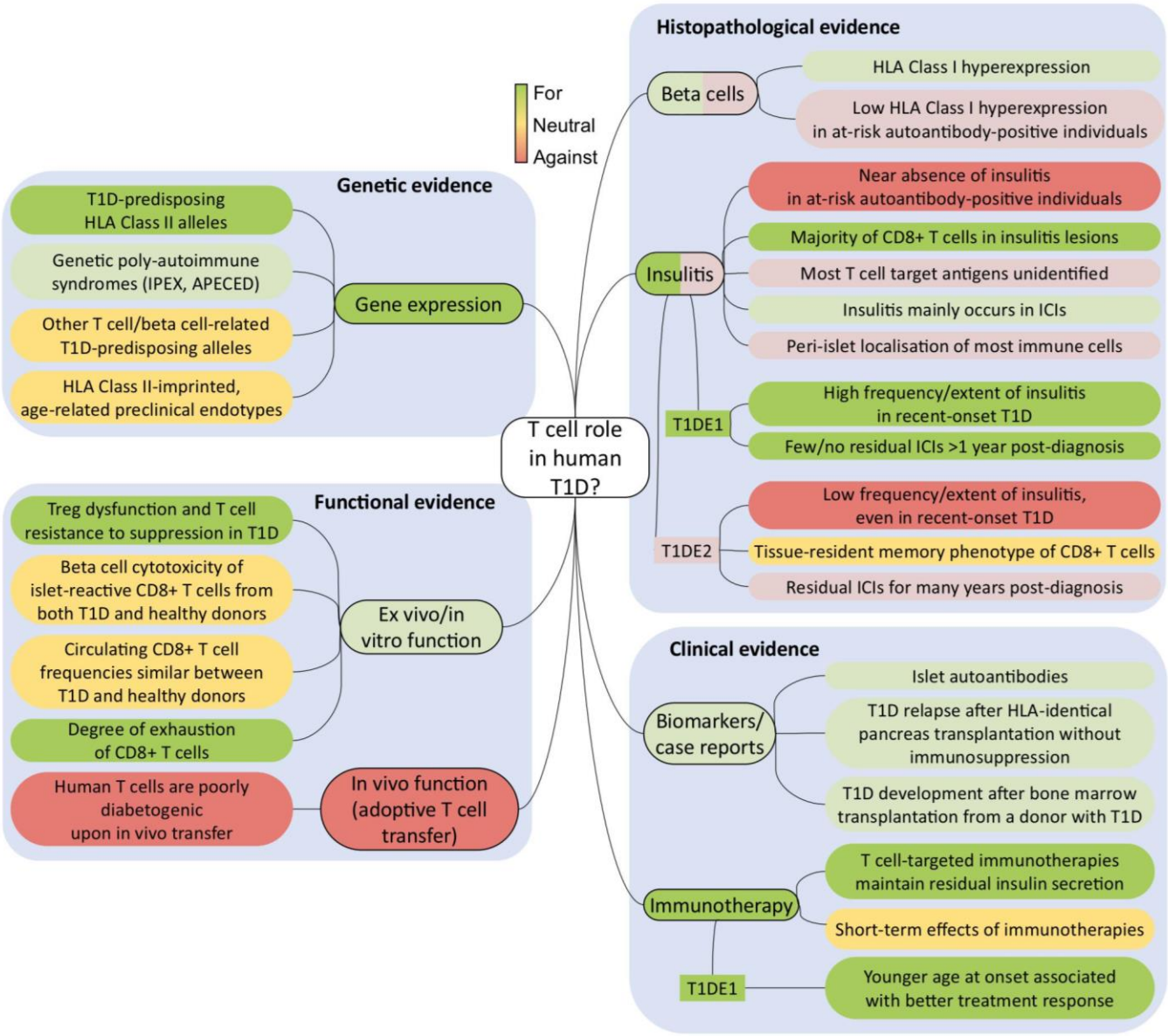


Fig. 1 Evidence for or against a primary pathogenic role for T cells in human type 1 diabetes (T1D). APECED, autoimmune polyendocrinopathy, candidiasis, ectodermal dystrophy (syndrome); IPEX, immune dysregulation, polyendocrinopathy, enteropathy, X-linked (syndrome); This figure is available as part of a downloadable slideset

and first seroconvert for anti-insulin autoantibodies (IAA), while older children (>40 months old) more often carry *HLA-DR3/DQ2* and present with anti-GAD at first seroconversion [5].

Against Although several disease-predisposing alleles are expressed in beta cells and modulate islet inflammation [1], their contribution to disease predisposition is rather small, suggesting a modulatory rather than driving role. Interestingly, HLA-II-imprinted endotypes influence seroconversion [5] but do not strongly affect age of onset [6], clinical progression or C-peptide loss post-diagnosis [7], indicating that other genetic and/or acquired factors determine the progression rate of beta cell damage.

Summing up: The genetics of type 1 diabetes suggests an involvement of T cells.

Histopathological evidence

For HLA-I hyperexpression on beta cells [8, 9] suggests a role for $CD8^+$ T cells. The second histopathological hallmark is the presence of immune infiltrates (insulitis) dominated by $CD8^+$ T cells, followed by $CD68^+$ macrophages, $CD20^+$ B cells and $CD4^+$ T cells [10–13]. Scattered neutrophils are also observed throughout the pancreas [14, 15] and evidence of dysregulated systemic innate immunity has been consistently reported, even before autoantibody seroconversion [14]. The relative under-representation of $CD4^+$ T cells and the near absence of regulatory T cells (Tregs) may reflect an earlier, more peripheral role (e.g. in pancreatic lymph nodes), while the over-representation of $CD8^+$ T cells may reflect a final effector role in beta cell destruction.

Studies on the Exeter Archival Diabetes Biobank (EADB) documented that insulitis of donors with recent-onset type 1 diabetes differs according to age [13, 16, 17], suggesting the existence of two endotypes with different pathogenic mechanisms (Fig. 2): type 1 diabetes endotype 1 (T1DE1), in individuals with onset at age <13 years; and type 1 diabetes endotype 2 (T1DE2), mostly occurring in individuals diagnosed at ≥13 years of age. T1DE1 is characterised by a $CD20^{high}$ insulitis rich in $CD8^+$ T cells, few residual insulin-containing islets (ICIs; indicative of residual beta cells) and evidence of abnormal insulin processing in the remaining beta cells. In contrast, T1DE2 displays a $CD20^{low}$ insulitis with fewer $CD8^+$ T cells, more residual ICIs and normal insulin processing. Recent-onset (<6 month) $CD20^{high}$ donors have a median of 78% of insulitis-positive ICIs (range 19–100%; $n = 20$), vs 14% (range 0–52%; $n = 19$) in $CD20^{low}$ donors (SJR, unpublished data). In line with this endotype view, the frequency of insulitis within 1 month of diagnosis is 73% in young donors with type 1 diabetes (<14 years old, likely to be mostly T1DE1) [18] and the parallel reduction in residual ICIs may indicate an aggressive beta cell destruction. In contrast, insulitis within 1 month from diagnosis is less frequent (29%) in older donors (15–40 years old, likely T1DE2) and long-term ICI preservation suggests reduced aggressiveness [19].

Against If viewing type 1 diabetes as a single disease, the histopathological evidence for a critical T cell involvement is unconvincing. Immune infiltrates are limited, as exemplified by the debate on the diagnostic definition of insulitis being the presence of ≥15 immune ($CD45^+$) cells/islet in ≥3 islets [20, 21]. This definition should be interpreted by considering the following points: (1) 15 immune cells/islet is only twice the number found in non-diabetic donors [10]; (2) it does not account for differences in islet size (i.e. the insulitis density); (3) infiltrated islets are typically few (<10%) and found in a minority (~20%) of donors [18, 22]; (4) peri-islet, non-invasive insulitis is more common [19]; and (5) lymphocytes are also found scattered in the exocrine tissue, even in autoantibody-positive non-diabetic donors [12]. Moreover, insulitis is mostly confined to ICIs (33% vs 2% in non-ICIs) and beta cell area and mass are higher in donors with type 1 diabetes with insulitis than in those without insulitis [22]. These patterns may indicate that immune cells leave the islets once beta cell destruction is complete.

These findings become more convincing when interpreted according to age-related endotype differences. However, a major bias is introduced by the specimens available, which are very limited for children with recent-onset type 1 diabetes because, fortunately, very few still die close to disease onset. For instance, the Network of Pancreatic Organ Donors with Diabetes (nPOD) has collected only one such specimen since 2007. In contrast, most tissues are available for T1DE2 donors with older-onset type 1 diabetes (i.e. those with weaker evidence for a major T cell role). In most studies, age-related endotype differences are further blurred by the analysis of specimens from individuals with overall long disease durations, which, together with age, influence the extent of insulitis and the number of residual ICIs. For instance, the frequency of insulitis in young donors with type 1 diabetes (<14 years old) drops from 73% within 1 month of diagnosis to 4% beyond 1 year [18]. Thus, even for the younger histopathological T1DE1 with stronger evidence for a major T cell role, key hallmarks have likely waned in most long-standing diabetes specimens available. Together with the focus of most studies on older-onset donors, these limitations may bias our view against T cells.

The bias of long disease duration is absent in the Diabetes Virus Detection Study (DiViD), which collected surgical pancreas specimens from living adults with new-onset disease

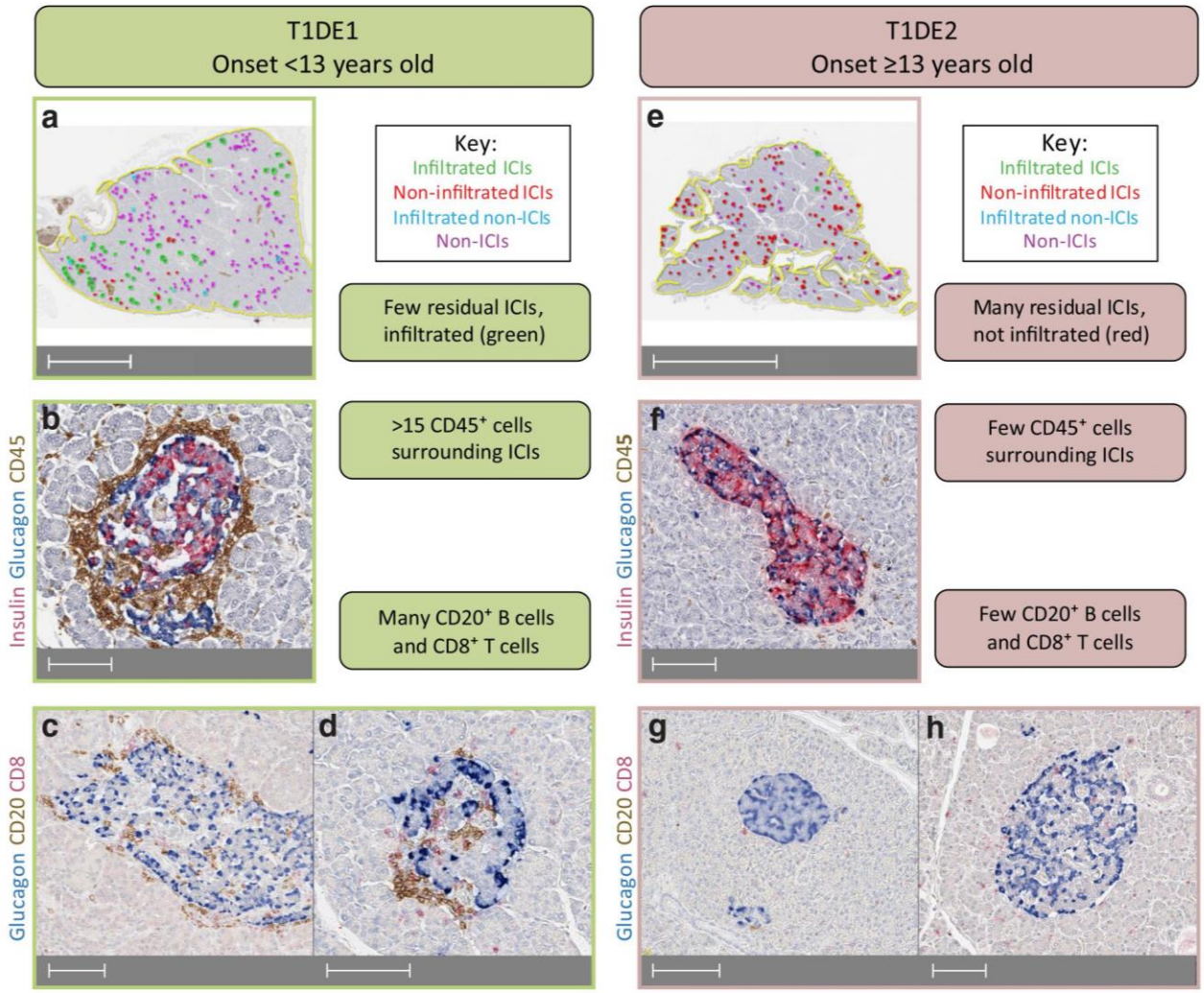


Fig. 2 Key differences in insulitis and residual ICIs in type 1 diabetes endotypes T1DE1 and T1DE2. (**a**–**d**) T1DE1 donors (mostly with a disease onset <13 years old) have few residual ICIs (**a**; green and red circles) compared with non-ICIs (blue and pink circles). The majority of residual ICIs display significant immune infiltration (**a**; >15 $CD45^+$ cells, green circles); a representative image is shown in (**b**). This infiltrate is enriched in $CD20^+$ B cells and $CD8^+$ T cells (**c**, **d**). (**e**–**h**) T1DE2 donors (mostly with a disease onset ≥13 years old) retain significant residual ICIs (**e**; red and green circles), the majority of which do not meet the criteria for insulitis (**e**; red circles). Most ICIs have limited infiltration with $CD45^+$ cells (**f**), the majority of which are $CD8^+$ T cells with few $CD20^+$ B cells (**g**, **h**). Specimens from representative donors from the EADB biobank (https://foulis.vub.ac.be) were immunostained with antibodies against insulin/glucagon/CD45 or glucagon/CD20/CD8 using triple chromogen-based immunohistochemistry similar to that described in [78]. Scale bars, 5 mm (**a**), 100 μm (**b**–**d**, **f**–**h**) or 10 mm (**e**). This figure is available as part of a downloadable slideset

[23]. In these donors, 11% of the islets displayed insulitis, a proportion comparable with that observed in similar age/duration-matched $CD20^{low}$ donors in both EADB and nPOD specimens. The fact that these samples came from living donors rules out perimortem or post-mortem changes as possible confounders. Collectively, the DiViD, EADB and nPOD findings with older-onset but short-duration disease rule out the possibility that the limited insulitis observed represents an extinguished phase of islet autoimmunity, as could be the case in donors with longer disease duration.

Even in individuals with recent-onset disease, the autoimmune process dates back several months or years and the situation could be different at the preclinical disease stages: stage 1, presence of ≥2 autoantibodies; stage 2, presence of autoantibodies and initial metabolic alterations (e.g. loss of first-phase insulin secretion); and stage 3, presence of autoantibodies and overt hyperglycaemia (i.e. clinical disease) [24]. One single autoantibody marks a low risk of progression, especially in adults, and is therefore not included in this staging system. Yet, most autoantibody-positive pancreas donors analysed to date are adults who are positive for a single autoantibody. It should be noted, however, that the low-risk definition of the single autoantibody status is based on longitudinal assessments in living humans and that the probability of appearance of a second autoantibody is approximately 7–20% at 5 years (depending on age), with a subsequent risk of clinical progression similar to those individuals who are positive for multiple autoantibodies from the start [25]. Thus, out of the reasonably

large series of ~100 donors positive for a single autoantibody analysed to date, a fraction would have eventually progressed. Yet, insulitis was invariably absent in all autoantibody-positive donors, even in those with predisposing HLA haplotypes [19]. Additionally, in the relatively limited number of specimens from donors positive for multiple autoantibodies (i.e. stage 1 disease; ~20 analysed to date), little or no insulitis was observed (1–9% of islets) [10, 22, 26]. This picture may reflect a disease kinetics remaining stable for many years before sudden progression. Arguably, snap-shooting this 'point of no-return' may be difficult in histopathological studies.

HLA-I hyperexpression is also typically found in ICIs; expression decreases with disease duration and is often associated with $CD8^+$ insulitis [27]. However, it is more readily apparent than insulitis and is present in most ICIs, probably reflecting a readout of beta cell stress. A possible interpretation of this discrepancy is that insulitis hallmarks may be more labile due to immune cells migrating between islets and the peri-islet and exocrine tissue, resulting in a variable pattern at any given time point compared with the more stable HLA-I hyperexpression. Nonetheless, even HLA-I hyperexpression was only found in 13% of islets from individuals positive for two autoantibodies, associated with higher yet mild $CD8^+$ T cell infiltration [28].

Another point of controversy is that, despite the identification of several antigen specificities among pancreas-infiltrating T cells, a large fraction remains unassigned. Moreover, whether these T cells present an effector/memory phenotype compatible with an active autoimmune engagement is unknown. One study suggested a prevalent (46%) tissue-resident memory phenotype of peri-islet $CD8^+$ T cells [29] that is enriched in non-diabetic donors [30] and, possibly, individuals who are T1DE2 [16, 29]. Although described as non-cytotoxic in insulitis lesions [29], tissue-resident T cells can proliferate locally, produce proinflammatory cytokines and recruit circulating cells, thus possibly contributing to perpetuate disease. Mouse studies [31, 32] suggest that a significant fraction of pancreas-infiltrating T cells are naive (also in humans [15]) and non-islet-reactive, possibly configuring insulitis as an 'open' lesion in which bystander T cell activation may contribute to beta cell damage.

Summing up: Most histopathological arguments suffer from sampling bias but are rather against a primary pathogenic role for T cells in T1DE2, while evidence is more substantial in T1DE1.

Functional evidence

For Although the NOD mouse model is imperfect, both $CD4^+$ and $CD8^+$ T cells are needed to transfer disease [33]. Moreover, NOD mice lacking MHC Class I (MHC-I) do not develop diabetes unless receiving splenocytes from diabetic animals [34]. Interestingly, beta cell-selective MHC-I knockout in mice [35] protects against diabetes but not insulitis, suggesting that MHC-I-interacting $CD8^+$ T cells exert a late pathogenic role, consistent with their final involvement in beta cell cytotoxicity. The antigen specificity of these T cells has been explored with T cell receptor (TCR)-transgenic NOD mouse studies, which suggest a variable pathogenic potency of different T cell clonotypes. Indeed, diabetes development is accelerated by some TCRs, such as the $CD4^+$ TCRs BDC-6.9 [36] (recognising a proinsulin/islet amyloid polypeptide hybrid peptide [37]) and 4.1 of unknown specificity [38] and the $CD8^+$ islet-specific glucose-6-phosphatase catalytic-subunit related protein $(IGRP)_{206–214}$-reactive 8.3 TCR [38]). Others do not accelerate diabetes development (i.e. the $CD4^+$ TCR BDC-2.5 reactive to a proinsulin/chromogranin-A hybrid peptide [37]), unless on a NOD/*scid* immunodeficient background [39], and the $CD8^+$ Ins_{B15-23}-reactive G9C8 TCR, which requires prior Ins_{B15-23} immunisation [40]).

Reports analysing Tregs concluded that, while the frequency of circulating forkhead box P3 $(FOXP3)^+CD4^+$ Tregs is unaltered in type 1 diabetes, their regulatory activity is diminished [41]. This alteration reflects both a reduced suppressive function of Tregs [42], sometimes associated with increased secretion of proinflammatory cytokines (IFN-γ, IL-17) [41, 43], and an increased resistance of conventional T cells to suppression [44]. Treg dysfunction is partly genetically imprinted (e.g. through *IL2RA* polymorphisms leading to unstable FOXP3 expression under limiting IL-2 concentrations [41]). It is also highly heterogeneous across individuals and largely overlapping with non-diabetic control individuals, possibly pointing to Treg-driven disease endotypes. Similarly, IL-10-polarised islet-reactive $CD4^+$ T cells were enriched in healthy donors and individuals with type 1 diabetes of later onset [45], who may be representative of T1DE2 [13].

Against Animal models of autoimmune diseases were initially produced by immunisation with organ extracts or antigens (e.g. for multiple sclerosis (1933) [46], orchitis, thyroiditis [47], adrenalitis, rheumatoid arthritis [48]). Despite a better knowledge of target antigens, similar immunisations of non-transgenic animals never provided an equivalent model of experimental insulitis/diabetes, possibly reflecting the requirement for beta cell dysfunction. The BioBreeding rat provides another spontaneous animal model of human type 1 diabetes. Interestingly, it harbours a profound systemic T cell lymphopenia, including near absence of $CD8^+$ T cells, which, despite some $CD8^+$ T cell infiltration in islets [49], is required for spontaneous diabetes development [50].

$CD8^+$ T cell clones recognising preproinsulin [51], IGRP [52] and zinc transporter 8 [53] epitopes can lyse beta cells in vitro. However, we documented that T cell clones from healthy donors display similar cytotoxic potency [53]. Although this similarity could reflect the long-term in vitro

stimulation of these clones, erasing ex vivo differences, it demonstrates that islet-reactive $CD8^+$ T cells from any individual can be differentiated into cytotoxic effectors. Albeit this does not negate a pathogenic role for T cells, it does not provide supportive evidence either, as frequently claimed. Moreover, most in vitro cytotoxic experiments [51, 52] employed high effector/target ratios (10/1 to 25/1) compared with those observed in situ in the pancreas [8].

These results mirror recent observations by us [53–55] and others [56–58] that islet-reactive $CD8^+$ T cells circulate at similar frequencies in healthy individuals and individuals with type 1 diabetes. It is unclear whether the same applies to $CD4^+$ T cells. These reports detected $CD8^+$ T cells using HLA-I multimers (i.e. independent of their functional phenotype), while a functional ELISpot readout of IFN-γ secretion repeatedly documented disease specificity [51, 53, 59, 60]. Thus, islet-reactive $CD8^+$ T cells may differ according to disease status in terms of functional profile (e.g. exhaustion) rather than frequency [57, 61] and in their capacity to home to the pancreas, where their density is enriched in donors with type 1 diabetes [53, 54]. This difference in homing may point to the heightened vulnerability of stressed/senescent beta cells as an important driver [1, 62, 63]. Overall, these results open key questions to better understand T cell autoimmunity and biomarkers but do not offer arguments either for or against a primary pathogenic role for T cells.

Human T cell adoptive transfer into immunodeficient mice was not diabetogenic (e.g. using an IGRP-reactive $CD8^+$ T cell clone [52]). Although TCR-transgenic $CD4^+$ T cells recognising the HLA-DQ8-restricted $INS_{B9\text{-}23}$ epitope transferred into HLA-DQ8-transgenic mice reconstituted with human thymus and $CD34^+$ cells induced diabetes [64], they required strong priming conditions (i.e. multiple low-dose streptozotocin and $INS_{B9\text{-}23}$ peptide immunisation). Streptozotocin-induced beta cell death and antigen release may provide the beta cell dysfunction signal needed for efficient T cell priming.

Summing up: Studies documenting Treg dysfunction and variable degrees of exhaustion of conventional T cells offer arguments for a role of human T cells. Arguments from ex vivo and in vitro human studies and in vivo 'humanised' mice are instead rather neutral or against such a role on its own and suggest an additional requirement for beta cell dysfunction.

Clinical evidence

For Islet autoantibodies appear before disease onset [5] and are produced by B cells, thus implying a helper role provided by $CD4^+$ T cells. Eight case reports documented disease relapse in non- or minimally immunosuppressed recipients with type 1 diabetes 6–12 weeks after pancreas transplantation from non-diabetic HLA-identical siblings, with no relapse observed under immunosuppression [65]. Conversely, a case report described disease development 4 years after non-T cell-depleted HLA-identical bone marrow transplantation from a donor with type 1 diabetes [66].

Among the panoply of immunotherapeutic agents trialled in individuals with new-onset type 1 diabetes, four have shown some effect on C-peptide preservation: LFA-3-Ig (alefacept) [67], CTLA-4-Ig (abatacept) [68], anti-CD20 (rituximab) [69] and anti-CD3 monoclonal antibodies [70–73]. Apart from the anti-B cell agent rituximab, all target mainly T cells, arguing for T cell-driven mechanisms targetable by drugs. Interestingly, superior clinical benefit is often observed in children [67, 69, 70, 72], who may represent the more T cell/B cell-driven endotype T1DE1. Concordantly, individuals diagnosed at younger ages generally lose C-peptide secretion more rapidly and the fast-progressing subgroup displays age-dependent blood gene expression and cell count profiles that are higher for B cells and lower for neutrophils [7]. Moreover, individuals with higher pre-treatment B cell counts achieve superior C-peptide preservation after rituximab treatment [7], exemplifying the relevance of disease endotypes to selection of more-targeted immunotherapies.

Against The limited and partial improvements observed in immunotherapy trials starkly contrast with the more significant benefits achieved in other autoimmune diseases. Examples of mainstay disease-modifying therapies include the following: IFN-β, glatiramer, dimethyl fumarate, sphingosine-1-phosphate receptor modulators and antibodies to very late antigen-4 and CD20 for multiple sclerosis [74]; and methotrexate, TNF and IL-6 inhibitors, rituximab and abatacept for rheumatoid arthritis [75]. However, all these agents perform much better when given early in the disease course, at a stage that might correspond to the preclinical phase of type 1 diabetes.

A recurrent observation from immunotherapy trials in individuals with new-onset diabetes is that the effect on C-peptide preservation is limited to the first months of treatment and resumes its decline thereafter [76]. While this may indicate a need for prolonged treatment, it could also reflect targeting of only the T cell pathogenic component, leaving others, possibly related to beta cells [1], free to drive further disease progression. However, trials with putative beta cell-protective agents (e.g. glucagon-like peptide-1 agonists) have so far failed to deliver significant clinical benefits, or to convincingly ascribe any such benefit to beta cell protection [77].

Summing up: Clinical evidence is mostly in favour of a pathogenic role for T cells and limited immunotherapeutic efficacy may be improved by endotype-based stratification and earlier treatment.

Technical limitations in the study of islet-reactive T cells that hamper definition of their pathogenic role and fulfilment of Koch's postulates

Incomplete knowledge of islet antigens

- Need for a comprehensive identification
- Need for a stringent degree of evidence (i.e. natural processing and presentation by beta cells/antigen-presenting cells and recognition by T cells)

Gaps to be addressed: Which islet-reactive T cells are pathogenic? Do they infiltrate the pancreas? Do target antigens differ according to age of onset?

Technologies to identify and isolate islet-reactive T cells

- Functional readouts (cytokine production, proliferation, upregulation of activation markers): more sensitive (lower T cell avidity threshold needed) but less specific (bystander activation of other T cells); alter the ex-vivo T cell phenotype; miss functionally silent T cells (e.g. anergic, exhausted); may miss Tregs
- 'Structural' readouts (binding of HLA-I/II multimers): less sensitive (higher T cell avidity threshold needed) but more specific; maintain the ex vivo T cell phenotype; detect T cells independently of their function, including Tregs
- Functional/structural readouts: may reduce the drawbacks of each

Gaps to be addressed: How does the frequency and/or function of circulating islet-reactive T cells differ between individuals with type 1 diabetes and healthy individuals? How do they differ according to age of onset?

Addressing Koch's postulate no. 1: Islet-reactive T cells must be present in all diseased individuals, but not in healthy ones. Currently: $CD8^+$ T cells are present in the blood of any individual but are enriched in the pancreas of individuals with type 1 diabetes; unclear for $CD4^+$ T cells

Addressing Koch's postulate no. 2: Islet-reactive T cells can be isolated from diseased individuals. Currently: Also from healthy individuals (both $CD4^+$ and $CD8^+$ T cells)

Different pathogenic/regulatory/homing potential?

Most studies are performed on peripheral blood

- Need to expand studies on pancreatic lymph nodes and islets
- Need to expand the scope of these studies: definition of antigen specificity (TCR sequencing/re-expression, in situ HLA-I/II multimer staining); more in-depth phenotyping (single-cell transcriptomics, laser-capture microdissection)
- Need assays to follow disease dynamics vs the static snapshot of histopathology: pancreas tissue slices, islet microfluidic and biosensor platforms

Gaps to be addressed: How do circulating and tissue T cells differ? Are all pancreas-infiltrating T cells reactive to islet antigens? Are they antigen-experienced, tissue-resident, cytotoxic? Can they migrate between islets, can they serially kill multiple beta cells over time?

T cell and beta cell functional studies are often performed in isolation

- Need for in vitro and in vivo models of T cell/beta cell crosstalk

Gaps to be addressed: How do beta cells modulate their autoimmune vulnerability? How do T cells modulate beta cell dysfunction? Can we identify clinically relevant beta cell-protective agents?

Addressing Koch's postulate no. 3: Type 1 diabetes must be induced when islet-reactive T cells are inoculated into a healthy host. Currently: One clinical case report; poor induction in humanised mouse models

Are multiple islet-reactive T cell specificities needed for efficient beta cell destruction? Initial beta cell damage/dysfunction required?

Addressing Koch's postulate no. 4: Islet-reactive T cells must be recoverable from inoculated hosts. Currently: Not attempted in the clinical case report; postulate fulfilled in humanised mouse models despite poor diabetes induction

Gaps in knowledge preventing conclusive evidence of the pathogenic role of T cells

The fulfilment of Koch's postulates for a pathogenic role for T cells is currently incomplete, partly due to technical limitations hindering progress (see Text box: 'Technical limitations in the study of islet-reactive T cells that hamper definition of their pathogenic role and fulfilment of Koch's postulates'). Conceptually, our understanding remains limited at two levels. First, it is not clear how environmental factors trigger T cell engagement. Most candidates (e.g. enteroviral infections and nutrients disrupting the gut barrier and/or microbiota composition) could exert their effects on both T cells and beta cells. Thus, the relative weight and temporal sequence of these effects would clarify whether T cell engagement represents a primary causative event or a bystander consequence of beta cell dysfunction, and whether such engagement affects disease initiation and/or progression/amplification. We also need to better understand the intermediate steps (e.g. metabolic and inflammatory derangements) connecting environmental triggers with T cell recruitment, and the innate-adaptive immune crosstalk. Second, differences in the frequency and/or phenotype of circulating islet-reactive T cells between diabetic and healthy donors are rather subtle if not absent altogether. We need to better understand the features of this universal state of 'benign' autoimmunity, how it is lost and its relationship with age. To this end, a detailed description of T cell modifications along the natural history of disease is required but remains largely unexplored in peripheral blood and even more so in the pancreas. Very few organ donors analysed are positive for multiple autoantibodies (stage 1 disease), and none are representative of stage 2 with initial dysglycaemia. We can speculate on whether insulitis may represent a transient phase in disease progression. Its near absence in stage 1 donors may thus reflect a relatively late autoimmune acceleration during stage 2 that we are presently missing.

Conclusion

Figure 1 summarises the available evidence for or against a primary pathogenic role for T cells in type 1 diabetes. While most of the evidence favours such a role, some does not provide the support often given as granted. Histopathological evidence yields a more conflicting picture. Although biased by some limitations inherent to these studies, they suggest the existence of two age-related endotypes. Clinical evidence of faster beta cell loss and superior benefit of immunotherapeutic intervention in individuals with younger-onset diabetes further support this endotype view. Several technical limitations and gaps in knowledge need to be filled to gauge T cell involvement according to age and disease endotypes. Type 1 diabetes may be a case of one name but two diseases (i.e. younger-onset with primary T cell-driven mechanisms and older-onset with primary beta cell-driven mechanisms) leading to similar clinical presentations but requiring different treatments.

Acknowledgements The authors would like to thank R. Wyatt (University of Exeter), I. Kusmartseva and M. Padilla (University of Florida) for the preparation and imaging of the pancreas sections.

Funding RM is funded by grants from The Leona M. and Harry B. Helmsley Charitable Trust (Helmsley no. 1901-03689), JDRF (2-SRA-2016-164-Q-R), the Fondation Francophone pour la Recherche sur le Diabète, the Fondation pour la Recherche Médicale (EQU20193007831) and the Agence Nationale de la Recherche (ANR-19-CE15-0014-01). SJR is funded by grants from Diabetes UK (16/0005480), JDRF (5-CDA-2014-221-A-N; 2-SRA-2018-474-S-B) and the MRC (MR/P010695/1). RM and SJR are funded and by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST). These Joint Undertakings receive support from the Union's Horizon 2020 research and innovation programme, European Federation of Pharmaceutical Industries Associations, JDRF and The Leona M. and Harry B. Helmsley Charitable Trust. Research in the laboratories of RM and SJR is performed with the support of the Network for Pancreatic Organ donors with Diabetes (nPOD; RRID:SCR_014641), a collaborative type 1 diabetes research project sponsored by JDRF (nPOD: 5-SRA-2018-557-Q-R) and The Leona M. & Harry B. Helmsley Charitable Trust (Grant#2018PG-T1D053). The content and views expressed are the responsibility of the authors and do not necessarily reflect the official view of nPOD. Organ Procurement Organizations (OPO) partnering with nPOD to provide research resources are listed at http://www.jdrfnpod.org/for-partners/npod-partners/.

Authors' relationships and activities The authors declare that there are no relationships or activities that might bias, or be perceived to bias, their work.

Contribution statement AC, SJR, EL and RM wrote the paper. All authors approved the version to be published.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Mallone R, Eizirik DL (2020) Presumption of innocence in type 1 diabetes: why are beta cells such vulnerable autoimmune targets? Diabetologia 63:1999–2006. https://doi.org/10.1007/s00125-020-05176-7
2. Battaglia M, Ahmed S, Anderson MS et al (2020) Introducing the endotype concept to address the challenge of disease heterogeneity

in type 1 diabetes. Diabetes Care 43(1):5–12. https://doi.org/10.2337/dc19-0880
3. Noble JA, Valdes AM, Varney MD et al (2010) HLA Class I and genetic susceptibility to type 1 diabetes. Diabetes 59(11):2972–2979. https://doi.org/10.2337/db10-0699
4. Robertson CC, Inshaw JRJ, Onengut-Gumuscu S, et al (2020) Fine-mapping, trans-ancestral and genomic analyses identify causal variants, cells, genes and drug targets for type 1 diabetes. bioRxiv 158071 (preprint). https://doi.org/10.1101/2020.06.19.158071
5. Pociot F, Lernmark Å (2016) Genetic risk factors for type 1 diabetes. Lancet 387(10035):2331–2339. https://doi.org/10.1016/S0140-6736(16)30582-7
6. Valdes AM, Erlich HA, Carlson J, Varney M, Moonsamy PV, Noble JA (2012) Use of class I and class II HLA loci for predicting age at onset of type 1 diabetes in multiple populations. Diabetologia 55(9):2394–2401. https://doi.org/10.1007/s00125-012-2608-z
7. Dufort MJ, Greenbaum CJ, Speake C, Linsley PS (2019) Cell type-specific immune phenotypes predict loss of insulin secretion in new-onset type 1 diabetes. JCI Insight 4(4):e125556. https://doi.org/10.1172/jci.insight.125556
8. Coppieters KT, Dotta F, Amirian N et al (2012) Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients. J Exp Med 209(1):51–60. https://doi.org/10.1084/jem.20111187
9. Richardson SJ, Rodriguez-Calvo T, Gerling IC et al (2016) Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. Diabetologia 59(11):2448–2458. https://doi.org/10.1007/s00125-016-4067-4
10. In't Veld P, Lievens D, De Grijse J et al (2007) Screening for insulitis in adult autoantibody-positive organ donors. Diabetes 56(9):2400–2404. https://doi.org/10.2337/db07-0416
11. Willcox A, Richardson SJ, Bone AJ, Foulis AK, Morgan NG (2009) Analysis of islet inflammation in human type 1 diabetes. Clin Exp Immunol 155(2):173–181. https://doi.org/10.1111/j.1365-2249.2008.03860.x
12. Rodriguez-Calvo T, Ekwall O, Amirian N, Zapardiel-Gonzalo J, von Herrath MG (2014) Increased immune cell infiltration of the exocrine pancreas: a possible contribution to the pathogenesis of type 1 diabetes. Diabetes 63(11):3880–3890. https://doi.org/10.2337/db14-0549
13. Arif S, Leete P, Nguyen V et al (2014) Blood and islet phenotypes indicate immunological heterogeneity in type 1 diabetes. Diabetes 63(11):3835–3845. https://doi.org/10.2337/db14-0365
14. Vecchio F, Buono NL, Stabilini A et al (2018) Abnormal neutrophil signature in the blood and pancreas of presymptomatic and symptomatic type 1 diabetes. JCI Insight 3(18):e122146. https://doi.org/10.1172/jci.insight.122146
15. Damond N, Engler S, Zanotelli VRT et al (2019) A map of human type 1 diabetes progression by imaging mass cytometry. Cell Metab 29(3):755–768.e5. https://doi.org/10.1016/j.cmet.2018.11.014
16. Leete P, Willcox A, Krogvold L et al (2016) Differential insulitic profiles determine the extent of β-cell destruction and the age at onset of type 1 diabetes. Diabetes 65(5):1362–1369. https://doi.org/10.2337/db15-1615
17. Leete P, Oram RA, McDonald TJ et al (2020) Studies of insulin and proinsulin in pancreas and serum support the existence of aetiopathological endotypes of type 1 diabetes associated with age at diagnosis. Diabetologia 63(6):1258–1267. https://doi.org/10.1007/s00125-020-05115-6
18. In't Veld P (2011) Insulitis in human type 1 diabetes. Islets 3(4):131–138. https://doi.org/10.4161/isl.3.4.15728
19. Rodriguez-Calvo T, Richardson SJ, Pugliese A (2018) Pancreas pathology during the natural history of type 1 diabetes. Curr Diab Rep 18(11):124. https://doi.org/10.1007/s11892-018-1084-3
20. Campbell-Thompson ML, Atkinson MA, Butler AE et al (2013) The diagnosis of insulitis in human type 1 diabetes. Diabetologia 56(11):2541–2543. https://doi.org/10.1007/s00125-013-3043-5
21. Donath MY, Hess C, Palmer E (2014) What is the role of autoimmunity in type 1 diabetes? A clinical perspective. Diabetologia 57(4):653–655. https://doi.org/10.1007/s00125-013-3153-0
22. Campbell-Thompson M, Fu A, Kaddis JS et al (2016) Insulitis and β-cell mass in the natural history of type 1 diabetes. Diabetes 65(3):719–731. https://doi.org/10.2337/db15-0779
23. Krogvold L, Wiberg A, Edwin B et al (2016) Insulitis and characterisation of infiltrating T cells in surgical pancreatic tail resections from patients at onset of type 1 diabetes. Diabetologia 59(3):492–501. https://doi.org/10.1007/s00125-015-3820-4
24. Insel RA, Dunne JL, Atkinson MA et al (2015) Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. Diabetes Care 38(10):1964–1974. https://doi.org/10.2337/dc15-1419
25. Bingley PJ, Boulware DC, Krischer JP, Type 1 Diabetes TrialNet Study Group (2016) The implications of autoantibodies to a single islet antigen in relatives with normal glucose tolerance: development of other autoantibodies and progression to type 1 diabetes. Diabetologia 59(3):542–549. https://doi.org/10.1007/s00125-015-3830-2
26. Wiberg A, Granstam A, Ingvast S et al (2015) Characterization of human organ donors testing positive for type 1 diabetes-associated autoantibodies. Clin Exp Immunol 182(3):278–288. https://doi.org/10.1111/cei.12698
27. Richardson SJ, Morgan NG, Foulis AK (2014) Pancreatic pathology in type 1 diabetes mellitus. Endocr Pathol 25(1):80–92. https://doi.org/10.1007/s12022-014-9297-8
28. Rodriguez-Calvo T, Suwandi JS, Amirian N et al (2015) Heterogeneity and lobularity of pancreatic pathology in type 1 diabetes during the prediabetic phase. J Histochem Cytochem 63(8):626–636. https://doi.org/10.1369/0022155415576543
29. Kuric E, Seiron P, Krogvold L et al (2017) Demonstration of tissue resident memory CD8 T cells in insulitic lesions in adult patients with recent-onset type 1 diabetes. Am J Pathol 187(3):581–588. https://doi.org/10.1016/j.ajpath.2016.11.002
30. Radenkovic M, Uvebrant K, Skog O et al (2017) Characterization of resident lymphocytes in human pancreatic islets. Clin Exp Immunol 187(3):418–427. https://doi.org/10.1111/cei.12892
31. Magnuson AM, Thurber GM, Kohler RH, Weissleder R, Mathis D, Benoist C (2015) Population dynamics of islet-infiltrating cells in autoimmune diabetes. Proc Natl Acad Sci U S A 112(5):1511–1516. https://doi.org/10.1073/pnas.1423769112
32. Christoffersson G, Chodaczek G, Ratliff SS, Coppieters K, von Herrath MG (2018) Suppression of diabetes by accumulation of non-islet-specific $CD8^+$ effector T cells in pancreatic islets. Sci Immunol 3(21):eaam6533. https://doi.org/10.1126/sciimmunol.aam6533
33. Bendelac A, Carnaud C, Boitard C, Bach JF (1987) Syngeneic transfer of autoimmune diabetes from diabetic NOD mice to healthy neonates. Requirement for both L3T4+ and Lyt-2+ T cells. J Exp Med 166(4):823–832. https://doi.org/10.1084/jem.166.4.823
34. Wicker LS, Leiter EH, Todd JA et al (1994) β2-Microglobulin-deficient NOD mice do not develop insulitis or diabetes. Diabetes 43(3):500–504. https://doi.org/10.2337/diab.43.3.500
35. Hamilton-Williams EE, Palmer SE, Charlton B, Slattery RM (2003) Beta cell MHC class I is a late requirement for diabetes. Proc Natl Acad Sci U S A 100(11):6688–6693. https://doi.org/10.1073/pnas.1131954100
36. Pauza ME, Dobbs CM, He J et al (2004) T-cell receptor transgenic response to an endogenous polymorphic autoantigen determines susceptibility to diabetes. Diabetes 53(4):978–988. https://doi.org/10.2337/diabetes.53.4.978

37. Delong T, Wiles TA, Baker RL et al (2016) Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. Science 351(6274):711–714. https://doi.org/10.1126/science.aad2791
38. Verdaguer J, Schmidt D, Amrani A, Anderson B, Averill N, Santamaria P (1997) Spontaneous autoimmune diabetes in monoclonal T cell nonobese diabetic mice. J Exp Med 186(10):1663–1676. https://doi.org/10.1084/jem.186.10.1663
39. Kurrer MO, Pakala SV, Hanson HL, Katz JD (1997) Beta cell apoptosis in T cell-mediated autoimmune diabetes. Proc Natl Acad Sci U S A 94(1):213–218. https://doi.org/10.1073/pnas.94.1.213
40. Wong FS, Khai Siew L, Scott G et al (2009) Activation of insulin-reactive CD8 T-cells for development of autoimmune diabetes. Diabetes 58(5):1156–1164. https://doi.org/10.2337/db08-0800
41. Hull CM, Peakman M, Tree TIM (2017) Regulatory T cell dysfunction in type 1 diabetes: what's broken and how can we fix it? Diabetologia 60(10):1839–1850. https://doi.org/10.1007/s00125-017-4377-1
42. Lindley S, Dayan CM, Bishop A, Roep BO, Peakman M, Tree TIM (2005) Defective suppressor function in $CD4^+CD25^+$ T-cells from patients with type 1 diabetes. Diabetes 54(1):92–99. https://doi.org/10.2337/diabetes.54.1.92
43. McClymont SA, Putnam AL, Lee MR et al (2011) Plasticity of human regulatory T cells in healthy subjects and patients with type 1 diabetes. J Immunol 186(7):3918–3926. https://doi.org/10.4049/jimmunol.1003099
44. Schneider A, Rieck M, Sanda S, Pihoker C, Greenbaum C, Buckner JH (2008) The effector T cells of diabetic subjects are resistant to regulation via $CD4^+FOXP3^+$ regulatory T cells. J Immunol 181(10):7350–7355. https://doi.org/10.4049/jimmunol.181.10.7350
45. Arif S, Tree TI, Astill TP et al (2004) Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health. J Clin Invest 113(3):451–463. https://doi.org/10.1172/JCI19585
46. Rivers TM, Sprunt DH, Berry GP (1933) Observations on attempts to produce acute disseminated encephalomyelitis in monkeys. J Exp Med 58(1):39–53. https://doi.org/10.1084/jem.58.1.39
47. Rose NR, Witebsky E (1956) Studies on organ specificity. V. Changes in the thyroid glands of rabbits following active immunization with rabbit thyroid extracts. J Immunol 76(6):417–427
48. Trentham DE, Townes AS, Kang AH (1977) Autoimmunity to type II collagen an experimental model of arthritis. J Exp Med 146(3):857–868. https://doi.org/10.1084/jem.146.3.857
49. Hosszufalusi N, Chan E, Teruya M, Takei S, Granger G, Charles MA (1993) Quantitative phenotypic and functional analyses of islet immune cells before and after diabetes onset in the BB rat. Diabetologia 36(11):1146–1154. https://doi.org/10.1007/BF00401059
50. Mordes JP, Bortell R, Blankenhorn EP, Rossini AA, Greiner DL (2004) Rat models of type 1 diabetes: Genetics, environment, and autoimmunity. ILAR J 45(3):278–291. https://doi.org/10.1093/ilar.45.3.278
51. Skowera A, Ellis RJ, Varela-Calviño R et al (2008) CTLs are targeted to kill β cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope. J Clin Invest 118(10):3390–3402. https://doi.org/10.1172/JCI35449
52. Unger WWJ, Pearson T, Abreu JRF et al (2012) Islet-specific CTL cloned from a type 1 diabetes patient cause beta-cell destruction after engraftment into HLA-A2 transgenic NOD/SCID/IL2RG null mice. PLoS One 7(11):e49213. https://doi.org/10.1371/journal.pone.0049213
53. Culina S, Lalanne AI, Afonso G et al (2018) Islet-reactive $CD8^+$ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 3(20):eaao4013. https://doi.org/10.1126/sciimmunol.aao4013
54. Gonzalez-Duque S, Azoury ME, Colli ML et al (2018) Conventional and neo-antigenic peptides presented by β cells are targeted by circulating naïve CD8+ T cells in type 1 diabetic and healthy donors. Cell Metab 28(6):946–960. https://doi.org/10.1016/j.cmet.2018.07.007
55. Azoury ME, Tarayrah M, Afonso G, et al (2020) Peptides Derived From Insulin Granule Proteins are Targeted by CD8+ T Cells Across MHC Class I Restrictions in Humans and NOD Mice. Diabetes 69(9). https://doi.org/10.2337/db20-0013
56. Skowera A, Ladell K, McLaren JE et al (2015) β-Cell-specific CD8 T cell phenotype in type 1 diabetes reflects chronic autoantigen exposure. Diabetes 64(3):916–925. https://doi.org/10.2337/db14-0332
57. Wiedeman AE, Muir VS, Rosasco MG et al (2019) Autoreactive CD8+ T cell exhaustion distinguishes subjects with slow type 1 diabetes progression. J Clin Invest 130(1):480–490. https://doi.org/10.1172/JCI126595
58. Yeo L, Pujol-Autonell I, Baptista R et al (2020) Circulating β cell-specific $CD8^+$ T cells restricted by high-risk HLA class I molecules show antigen experience in children with and at risk of type 1 diabetes. Clin Exp Immunol 199(3):263–277. https://doi.org/10.1111/cei.13391
59. Mallone R, Martinuzzi E, Blancou P et al (2007) $CD8^+$ T-cell responses identify β-cell autoimmunity in human type 1 diabetes. Diabetes 56(3):613–621. https://doi.org/10.2337/db06-1419
60. Scotto M, Afonso G, Larger E et al (2012) Zinc transporter $(ZnT)8_{186\text{-}194}$ is an immunodominant $CD8^+$ T cell epitope in HLA-$A2^+$ type 1 diabetic patients. Diabetologia 55(7):2026–2031. https://doi.org/10.1007/s00125-012-2543-z
61. McKinney EF, Lee JC, Jayne DRW, Lyons PA, Smith KGC (2015) T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. Nature 523(7562):612–616. https://doi.org/10.1038/nature14468
62. Thompson PJ, Shah A, Ntranos V, Van Gool F, Atkinson M, Bhushan A (2019) Targeted elimination of senescent beta cells prevents type 1 diabetes. Cell Metab 29(5):1045–1060.e10. https://doi.org/10.1016/j.cmet.2019.01.021
63. Lee H, Lee Y-S, Harenda Q, et al (2020) Beta cell dedifferentiation induced by IRE1α deletion prevents type 1 diabetes. Cell Metab 31(4):822–836.e5. https://doi.org/10.1016/j.cmet.2020.03.002
64. Tan S, Li Y, Xia J et al (2017) Type 1 diabetes induction in humanized mice. Proc Natl Acad Sci U S A 114(41):10954–10959. https://doi.org/10.1073/pnas.1710415114
65. Sutherland DE, Goetz FC, Sibley RK (1989) Recurrence of disease in pancreas transplants. Diabetes 38(S1):85–87. https://doi.org/10.2337/diab.38.1.s85
66. Lampeter EF, Homberg M, Quabeck K et al (1993) Transfer of insulin-dependent diabetes between HLA-identical siblings by bone marrow transplantation. Lancet 341(8855):1243–1244. https://doi.org/10.1016/0140-6736(93)91148-f
67. Rigby MR, Harris KM, Pinckney A et al (2016) Alefacept provides sustained clinical and immunological effects in new-onset type 1 diabetes patients. J Clin Invest 125(8):3285–3296. https://doi.org/10.1172/JCI81722
68. Orban T, Bundy B, Becker DJ et al (2014) Costimulation modulation with abatacept in patients with recent-onset type 1 diabetes: follow-up 1 year after cessation of treatment. Diabetes Care 37(4):1069–1075. https://doi.org/10.2337/dc13-0604
69. Pescovitz MD, Greenbaum CJ, Krause-Steinrauf H et al (2009) Rituximab, B-lymphocyte depletion, and preservation of beta-cell function. N Engl J Med 361(22):2143–2152. https://doi.org/10.1056/NEJMoa0904452
70. Keymeulen B, Walter M, Mathieu C et al (2010) Four-year metabolic outcome of a randomised controlled CD3-antibody trial in

recent-onset type 1 diabetic patients depends on their age and baseline residual beta cell mass. Diabetologia 53(4):614–623. https://doi.org/10.1007/s00125-009-1644-9

71. Aronson R, Gottlieb PA, Christiansen JS et al (2014) Low-dose otelixizumab anti-CD3 monoclonal antibody DEFEND-1 study: results of the randomized phase iii study in recent-onset human type 1 diabetes. Diabetes Care 37(10):2746–2754. https://doi.org/10.2337/dc13-0327
72. Sherry N, Hagopian W, Ludvigsson J et al (2011) Teplizumab for treatment of type 1 diabetes (Protégé study): 1-year results from a randomised, placebo-controlled trial. Lancet 378(9790):487–497. https://doi.org/10.1016/S0140-6736(11)60931-8
73. Herold KC, Bundy BN, Long SA et al (2019) An anti-CD3 antibody, teplizumab, in relatives at risk for type 1 diabetes. N Engl J Med 381(7):603–613. https://doi.org/10.1056/NEJMoa1902226
74. Cree BAC, Mares J, Hartung H-P (2019) Current therapeutic landscape in multiple sclerosis: an evolving treatment paradigm. Curr Opin Neurol 32(3):365–377. https://doi.org/10.1097/WCO.0000000000000700
75. Giacomelli R, Afeltra A, Alunno A et al (2017) International consensus: what else can we do to improve diagnosis and therapeutic strategies in patients affected by autoimmune rheumatic diseases (rheumatoid arthritis, spondyloarthritides, systemic sclerosis, systemic lupus erythematosus, antiphospholipid syndrome and Sjogren's syndrome)? The unmet needs and the clinical grey zone in autoimmune disease management. Autoimmun Rev 16(9):911–924. https://doi.org/10.1016/j.autrev.2017.07.012
76. Greenbaum CJ, Beam CA, Boulware D et al (2012) Fall in C-peptide during first 2 years from diagnosis: evidence of at least two distinct phases from composite type 1 diabetes TrialNet data. Diabetes 61(8):2066–2073. https://doi.org/10.2337/db11-1538
77. Kielgast U, Holst JJ, Madsbad S (2011) Antidiabetic actions of endogenous and exogenous GLP-1 in type 1 diabetic patients with and without residual β-cell function. Diabetes 60(5):1599–1607. https://doi.org/10.2337/db10-1790
78. Kusmartseva I, Beery M, Hiller H et al (2020) Temporal analysis of amylase expression in control, autoantibody-positive, and type 1 diabetes pancreatic tissues. Diabetes 69(1):60–66. https://doi.org/10.2337/db19-0554

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Annex 3: Review on β-cell-specific antigen processing and presentation pathways and implications in T1D

frontiers in Immunology

MINI REVIEW
published: 29 March 2021
doi: 10.3389/fimmu.2021.639682

Check for updates

Making Insulin and Staying Out of Autoimmune Trouble: The Beta-Cell Conundrum

Alexia Carré[1*] *and Roberto Mallone*[1,2*]

[1] *Université de Paris, Institut Cochin, CNRS, INSERM, Paris, France,* [2] *Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, Paris, France*

Autoimmune type 1 diabetes (T1D) results from the intricate crosstalk of various immune cell types. CD8+ T cells dominate the pro-inflammatory milieu of islet infiltration (insulitis), and are considered as key effectors of beta-cell destruction, through the recognition of MHC Class I-peptide complexes. The pathways generating MHC Class I-restricted antigens in beta cells are poorly documented. Given their specialized insulin secretory function, the associated granule processing and degradation pathways, basal endoplasmic reticulum stress and susceptibility to additional stressors, alternative antigen processing and presentation (APP) pathways are likely to play a significant role in the generation of the beta-cell immunopeptidome. As direct evidence is missing, we here intersect the specificities of beta-cell function and the literature about APP in other cellular models to generate some hypotheses on APPs relevant to beta cells. We further elaborate on the potential role of these pathways in T1D pathogenesis, based on the current knowledge of antigens presented by beta cells. A better understanding of these pathways may pinpoint novel mechanisms amenable to therapeutic targeting to modulate the immunogenicity of beta cells.

Keywords: antigen presentation, antigen processing, autophagy, crinophagy, insulin granule, MHC class I, neo-epitopes, signal peptide

OPEN ACCESS

Edited by:
Iñaki Alvarez,
Universitat Autònoma de Barcelona, Spain

Reviewed by:
Howard Davidson,
University of Colorado Anschutz Medical Campus, United States
Bruce Verchere,
University of British Columbia, Canada

***Correspondence:**
Alexia Carré
alexia.carre@inserm.fr
Roberto Mallone
roberto.mallone@inserm.fr

Specialty section:
This article was submitted to Antigen Presenting Cell Biology, a section of the journal Frontiers in Immunology

Received: *09 December 2020*
Accepted: *12 March 2021*
Published: *29 March 2021*

Citation:
Carré A and Mallone R (2021) Making Insulin and Staying Out of Autoimmune Trouble: The Beta-Cell Conundrum. Front. Immunol. 12:639682. doi: 10.3389/fimmu.2021.639682

INTRODUCTION

Type 1 diabetes (T1D) is an autoimmune disease characterized by the destruction of insulin-producing beta cells. It stems from a complex interplay of innate and adaptive immune cells. CD8+ T cells dominate the immune infiltration of islets and play a prominent role as final effectors of beta-cell loss (1). There is also growing evidence supporting the idea that beta-cell dysfunction is another key driver of T1D pathogenesis (2). The heterogeneity of pancreas histopathology between T1D donors and even across islets from the same pancreas, both in terms of immune infiltrates and residual beta cells, have led to the definition of age-related endotypes (3), in which the component of beta-cell dysfunction may be dominant in adult-onset cases (4, 5). Effector CD8+ T cells recognize MHC Class I (MHC-I)-peptide complexes at the surface of beta cells. The conventional MHC-I antigen processing and presentation (APP) machinery is a multi-step process (**Figure 1A**) where *i)* cytosolic proteins of microbial or self-origin are degraded into peptides in a proteasome-mediated manner; *ii)* the resulting peptides are transported into the endoplasmic reticulum (ER) *via* the

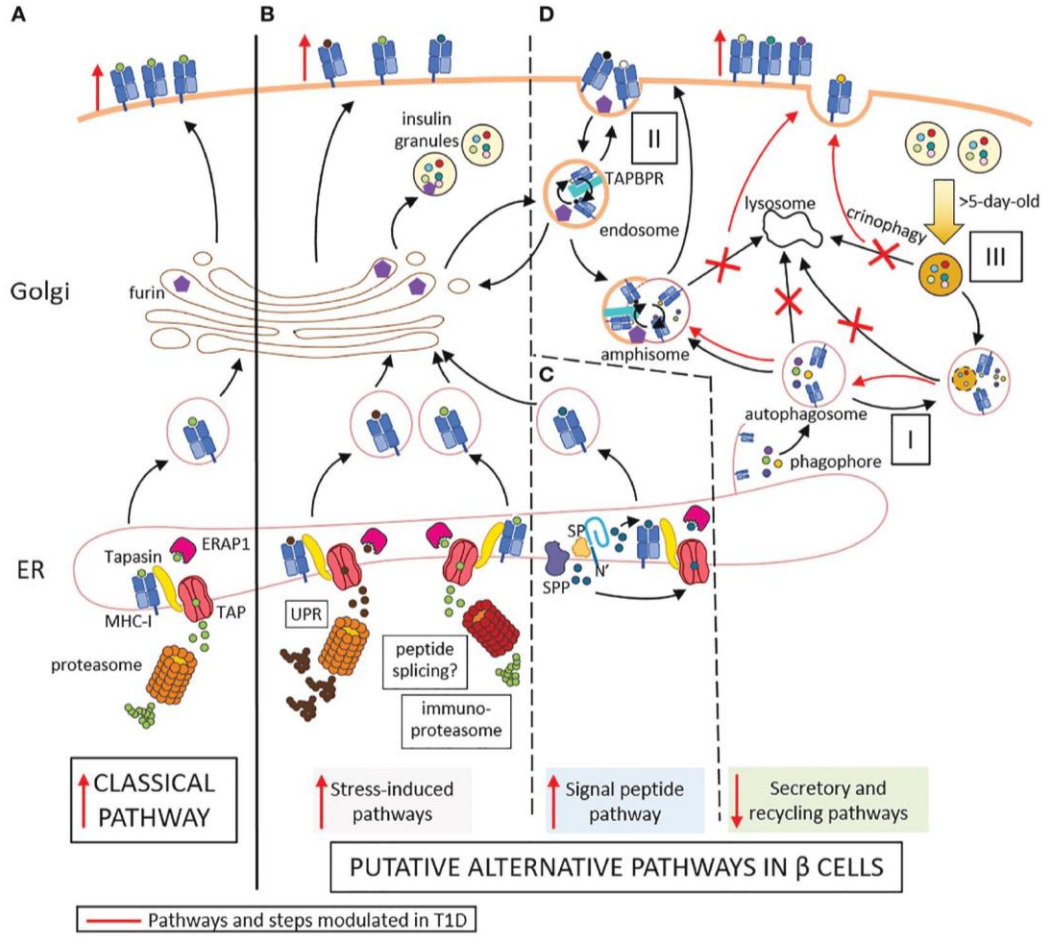


FIGURE 1 | Conventional **(A)** and putative alternative APP pathways **(B–D)** in beta cells. **(A)** In conventional APP pathways, misfolded proteins are degraded by the proteasome into peptides that are transported through TAP into the ER lumen. Here, they are further trimmed by ERAP1 and loaded onto MHC-I molecules that are associated to TAP *via* tapasin. The peptide-MHC-I complexes are exported to the Golgi and then to the cell surface. **(B)** Under inflammatory conditions, ER stress induces the UPR, which increases protein degradation and antigen presentation. The immuno-proteasome is also induced, further increasing APP and possibly resulting in alternative peptide splicing events and neo-epitopes generated. **(C)** Signal peptides are docked within the ER membrane. While the protein translation continues, the signal peptide is cleaved by the signal peptidase (SP) and further trimmed by the signal peptide peptidase (SPP). The N-terminal cleavage products of the signal peptide are released in the cytosol and can re-enter the ER through TAP for antigen presentation. The C-terminal cleavage products remaining in the ER can be loaded onto MHC-I molecules in a TAP- and proteasome-independent manner. **(D)** I. Cellular components are engulfed by nascent phagophore at the ER membrane, thus forming autophagosome, likely containing MHC-I molecules. II. Endocytosis of cell surface component forms endosomes, which can contain both MHC-I molecules and TAPBR promoting peptide exchange within endosomes. These vesicles can be directed to the Golgi apparatus, where peptides (free or released from MHC-I) can be trimmed by Golgi enzymes such as furins prior to loading onto MHC-I molecules. Autophagosomes, if not directly fused to lysosomes for degradation of their content, can fuse with endosome to form amphisomes, thus opening another peptide exchange pathway. These vesicles can subsequently be directed to lysosomes for degradation. III. Old insulin granules can be degraded by crinophagy. T1D induces changes (in red) characterized by enhanced ER stress **(B)**, increased insulin production and signal peptide processing **(C)**, and altered vesicle trafficking **(D)**, thus enhancing surface exposure of MHC-I complexes. Red marks indicate changes occuring in T1D: vertical arrows represent the increase or decrease of a step or pathway; bended arrows represent favored events; crosses represent blocked events. Some cellular components were modified from Servier Medical Art (smart.servier.com).

transporter associated with antigen processing (TAP). Here, *iii)* they are further trimmed by ER aminopeptidase (ERAP)1 prior to *iv)* loading on nascently-formed MHC-I molecules that are associated to TAP through tapasin; *v)* the stable peptide-MHC-I complexes are finally translocated to the Golgi complex and to the cell surface (6). This review addresses the specificities of the direct alternative MHC-I APP (aAPP) machinery within beta cells and its potential role in T1D pathogenesis (**Figures 1B–D**). It therefore focuses on the CD8+ T-cell responses that are triggered by MHC-I APP, notwithstanding the role of other cell types, and particularly of MHC-I and MHC-II APP by professional antigen-presenting cells (APCs) for the priming of naïve T cells. Since the beta-cell APP pathways are poorly if at all documented, we here provide a hypothesis-generating review that bridges the separate bodies of knowledge available for MHC-I APP and beta-cell biology.

SECRETING INSULIN: THE ACHILLES' HEEL OF BETA CELLS?

Like other endocrine cells, beta cells are hormone-secreting cells organized in glands that are highly vascularized, allowing the secretion of their products directly into the bloodstream. These features may influence the APP pathways used by these cells (2). Indeed, beta cells are equipped to sense changes in blood glucose levels and to respond by releasing appropriate amounts of insulin. For this purpose, they need to constantly adapt their secretory response to changes in metabolic and nutritional state. This adaptation is critical compared to that of other tissues because beta cells are long-lived and virtually non-proliferating, hence they can only modulate their function but not their cell numbers. Beta cells accommodate these metabolic variations by intensifying the synthesis of insulin and its precursor proinsulin. Under these intense bio-synthetic rates, misfolded proteins are more likely to accumulate within the ER, activating the unfolded protein response (UPR) (7). The UPR is a natural adaptive response that maintains cellular homeostasis by fulfilling three main tasks: to decrease the protein translation rate, to enhance the synthesis of protein-folding chaperones, and to increase the degradation of misfolded proteins into peptides through the proteasome (7). While part of the free cytosolic peptides generated by this process is further hydrolyzed to single amino acids, another portion is transported to the ER *via* TAP for the loading onto MHC-I molecules, thus shaping the catalog of MHC-I-bound peptides (immunopeptidome) presented at the beta-cell surface (**Figure 1B**). Beta cells rely on a well-balanced UPR to ensure their function and survival (8). Thus, beta cells are constantly functioning on the edge between physiological and pathological UPR and are exquisitely sensitive to additional stressors that can flip this delicate balance and lead to a compensatory maladaptive UPR further increasing ER stress prior to apoptosis (9, 10). Metabolic overload can act as an accelerator but is unlikely to drive this UPR transition on its own, as such overload is more common in type 2 diabetes and yet does not usually trigger islet autoimmunity. The environmental factors at play are likely multiple, but their identity and interaction with a susceptible genetic background remain ill-defined, with the possible exception of enteroviral infections (11). Considering the link between the UPR and APP through increased proteasomal degradation of misfolded proteins, the shift from physiological to pathological UPR may modulate the immunopeptidome and the beta-cell visibility to the immune system.

Under physiological conditions, the beta cells maintain consistent intracellular stores of insulin granules to allow their rapid exocytosis when the glycemia rises. While insulin granules are found by thousands in individual beta cells (12), not all are secreted. Indeed, the remarkable functional plasticity of beta cells includes recycling pathways to dispose of old (>5 days) or excess secretory granules. Commonly, insulin granules are degraded through macro-autophagy, by which they are engulfed by autophagosomes prior to degradation in lysosomes. Another endocrine-specific recycling pathway is a specialized macro-autophagy process named crinophagy, by which insulin granules fuse directly with lysosomes, thus generating so-called crinosomes (**Figure 1D**). While the role of these different types of autophagy has been extensively studied in type 2 diabetes, investigations in T1D are scanty. One recent study (13) reported defective macro-autophagy and diminished crinophagy in the beta cells of non-obese diabetic (NOD) mice compared to non-diabetic NOD littermates and non-obese resistant (NOR) mice. These conclusions were confirmed in T1D pancreas specimens and were based on the lack of co-localization of the lysosomal marker Lamp1 with either the granule marker proinsulin or the autophagosome marker LC3, thus suggesting a defect in the late stage of autophagy, when granules/autophagosomes fuse with lysosomes. An alternative explanation is that this reduced crinophagy may simply reflect a limited accumulation of old granules due to a compensatory higher insulin exocytosis rate in remaining beta cells. Interestingly, this study also reported a parallel increase of autophagosomes in insulin-positive cells of pancreata from autoantibody-positive non-diabetic donors in comparison to both control and T1D donors, whose autophagosome content was instead similar. As autophagy helps to prevent oxidative damage and promote cell survival, we can hypothesize that the beta-cell stress imposed by the inflammatory milieu of insulitis triggers an increase in their autophagy rate (14), prior to, and likely independently of, overt hyperglycemia. Interestingly, lysosomal autophagy can also degrade mitochondria (15). This process, known as mitophagy, is also part of stress responses (16), and is regulated by CLEC16A, whose genetic locus encodes variants associated with T1D risk (17). Pancreatic CLEC16A knockout mice harbor a defect in both mitochondrial and granule turnover (16), pointing to an intersection between the mitophagy and crinophagy pathway. It is thus possible that the altered granule disposal observed in stressed beta cells may also reflect overloading of this pathway by mitochondrial substrates. Its modulation by the *CLEC16A* genetic background may exacerbate the effect of stress on aAPPs. This feature also emphasizes the current view of lysosomes as dynamic structures that interact with different intracellular organelles and are continuously consumed and re-formed rather than just being end-stage degradation hubs (15).

Given this notion of impaired autophagy in T1D (13), what is the fate of the autophagosomes and insulin granules that are not degraded by lysosomes? In non-professional APCs, surface peptide-loaded MHC-I molecules seem to be recycled by clathrin-independent mechanisms, forming endosomes (18, 19). Whether and how these endosomes are degraded by lysosomes or targeted toward an early or late recycling pathway is unclear. Nonetheless, peptide exchange for MHC-I loading is likely to occur within the endosome. The TAP binding protein related (TAPBPR) chaperone is currently held as the main candidate for this role (20), and its preference for a slightly acidic to neutral pH is in the range of that found in endosomes. Another possibility is the encounter between endosomes and autophagosomes prior to recycling (21), forming so-called amphisomes (**Figure 1D**). Hence, the autophagosomes that are not degraded by lysosomes may be more likely to fuse with

endosomes and gain access to the cell surface. Considering that the autophagosome membrane is thought to derive from the ER and can thus harbor MHC-I molecules amenable to peptide loading (22), this process may provide an aAPP pathway. Similarly, insulin granules that are not degraded by lysosomes could indirectly participate in APP through engulfment by autophagosomes. Collectively, these data suggest that alternative autophagy-derived pathways for MHC-I loading could be operational in beta cells and participate in the recycling processes needed for disposing of insulin granules. While these pathways are active in resting beta cells, they may be further enriched in T1D as autophagy is impaired, possibly providing novel disease-enhanced pathways.

Besides insulin, secretory granules contain other proteins that we reported to dominate the beta-cell-specific immunopeptidome targeted by islet autoimmunity. First, a large fraction of MHC-I-bound peptides recovered from a human beta-cell line and primary islets originate from these granule proteins (23), which is not unexpected given the abundance and fast turnover of these proteins in beta cells (24). Second and most important however, these granule-derived peptides were also prominent targets recognized by circulating CD8+ T cells (23, 25). More recently, we found that H2-Kd-restricted peptides derived from the murine orthologues of some of these proteins, namely secretogranin-5, urocortin-3 and proconvertase-2, are also recognized by islet-infiltrating CD8 + T cells in prediabetic NOD mice (25). Moreover, CD8+ T cells recognizing these peptides were diabetogenic upon in-vivo transfer into NOD/*scid* recipients (25). Besides their localization in granules, it is noteworthy that these novel antigens (together with chromogranin A and others) (26) share several other features with insulin. First, they are all soluble proteins that are released along with insulin during granule exocytosis, a feature that could endow beta cells with a unique capability of sensitizing T cells at distance following APP of these proteins by extra-pancreatic APCs (2, 27). Second, they are all produced as precursors (pro-proteins) and subsequently undergo intermediate processing, first in the ER to cleave the signal peptide (see below), and then in immature granules through proconvertases, carboxypeptidase E (CPE) and furins that lead to their bioactive products (**Figure 2**). Intriguingly, reduced levels of proconvertases and CPE along with impaired proconvertase activity and proinsulin processing in beta cells has been repeatedly reported for T1D, even before clinical onset (4, 28–32). These enzymatic defects could lead to increase protein misfolding, which might feed the proteasomal recycling pathway and further fuel the APP machinery. It is also possible that granule protein byproducts may accumulate in the ER, the Golgi or the cytosol and thus become accessible for processing by other enzymes and for loading onto MHC-I molecules, either directly, through retrograde transport to the ER, or indirectly, *via* TAP. Of further note, proconvertases themselves as well as (pro) cathepsins are also synthesized as inactive precursors and subsequently activated, often through auto-enzymatic reactions. We can thus speculate that ER stress and impaired proconvertase activity may also negatively impact these processes since their earliest steps of ER export and self-activation.

An unsolved conundrum is how the T-cell targeting of granule antigens that are mostly shared with other endocrine cells such as alpha cells leads to an autoimmune response that is exquisitely beta-cell-specific. This question is even more

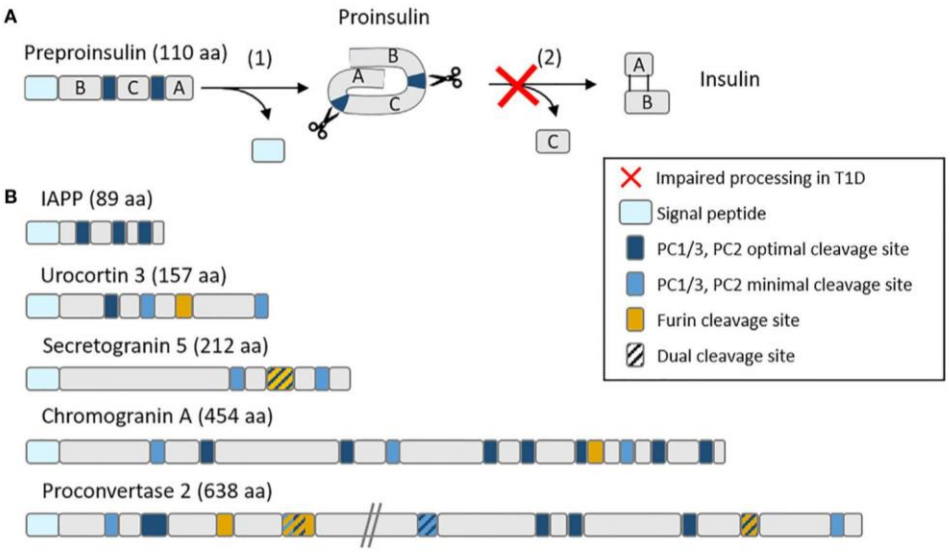


FIGURE 2 | Schematic representation of **(A)** the insulin maturation process and **(B)** mapping of proconvertase cleavage sites within insulin granule proteins. **(A)** Insulin is translated as a preprohormone comprising a signal peptide (light blue), followed by the B chain, C peptide and A chain. The first processing step (1) consists in the cleavage of the signal peptide, followed by (2) the cleavage by proconvertases of the C peptide and formation of disulfide bonds between the A and B chain. In T1D, this second step is often impaired, leading to proinsulin accumulation. **(B)** PC1/3 and PC2 optimal (RR, KR) and minimal (KK, RK) cleavage sites along with furin cleavage sites (R-X-X-R) (where X is any amino acid) are mapped on major insulin granules proteins. aa, amino acids.

important in light of recent reports suggesting that MHC–I upregulation is more prominent in alpha cells, even before clinical T1D (33). Protective factors at play in alpha cells may include their lower biosynthetic rate, the narrower dynamic range of their secretory response (34), their lower susceptibility to ER stress-induced apoptosis (35), and their higher expression of non-classical, inhibitory MHC-I molecules, i.e. human leukocyte antigen (HLA)-E (36). Moreover, autophagy seems less important for the maintenance of alpha-cell metabolism (37) than it is for beta cells (38), possibly making autophagy-related APP pathways less active.

NEO-EPITOPE GENERATION IN BETA CELLS

The large stores of insulin granules and the functional adaptability to metabolic changes of beta cells point to a remarkably active translation machinery, which enhances the odds of neo-epitope generation. Neo-epitopes are unconventional peptides such as post-translational modified sequences, defective ribosomal products (DRiPs), alternative mRNA splicing products and proteasomal peptide splicing products (i.e. resulting from the fusion of two non-adjacent regions of a protein). We reported that some of these products contribute to the MHC-I peptidome of beta cells (23, 25, 26). Very little is however known about how such peptides are generated and whether the required pathways are active in beta cells and/or professional APCs. In spite of their tissue specificity, the overall quantitative contribution of neo-epitopes to the beta-cell immunopeptidome appears to be minor, as only a limited number of MHC-I-restricted neo-epitopes has been reported to date (26). A possible scenario is that these neo-epitopes might be symptomatic of early beta-cell dysfunction, e.g. altered granule bio-synthesis/disposal, and become visible to T cells *via* the uptake of granules or of their exocytosed content by APCs (2). The combination of generation mechanisms favored by early beta-cell dysfunction and of recognition mechanisms favored by T cells that have escaped thymic deletion may endow them with a crucial role in the onset of islet autoimmunity.

Another aspect is the central role that the constitutive proteasome, and its inflammation-induced counterpart, the immunoproteasome, play in the generation of some of these neo-epitopes (39). It is noteworthy that in-vitro exposure of beta cells to interferon (IFN)-γ induces the expression of the immunoproteasome (40) which is known to enhance proteolysis, whereas immunoproteasome-deficient cells accumulate reactive oxygen species (ROS) (41). The immunoproteasome deficiency has also been investigated by in-vivo knockout in mice (42), which led to a CD8+ T-cell-mediated multi-organ autoimmunity comprising insulitis and CD8+ T cells reactive to an IGRP beta-cell peptide with low affinity H-2Kb binding. The authors hypothesized that the constitutive proteasome alone could fail at generating enough high-affinity MHC-I binders. This possibility is supported by other studies highlighting that the immunoproteasome uses the same cleavage sites as the constitutive proteasome, hence impacting the relative quantity rather than the quality of the peptides produced (39, 43, 44). Defective immunoproteasome activity and the resulting decrease in peptide output could thus allow weak-binding immunogenic peptides usually outcompeted by stronger binders to be processed and presented. Conversely, the inflammatory insulitis microenvironment of T1D may induce the well-known MHC-I upregulation (45), but also immunoproteasome expression, thus further enriching the immunopeptidome and favoring strong MHC binders. The higher binding affinity of these peptides may synergize with MHC-I upregulation to increase their availability for T-cell recognition. This increase in beta-cell visibility to T cells may provide one explanation to our recent observation that beta-cell-reactive CD8+ T cells circulate at similar frequencies in T1D and healthy donors (23, 46). This observation suggests that the difference between the 'benign' autoimmunity of healthy donors and the progressive autoimmunity of T1D patients may lie not only in their T-cell repertoire, but also in their beta-cell vulnerability (2).

PATHWAYS FOR SIGNAL PEPTIDE EPITOPES

Insulin is a protein hormone that is initially synthesized as a preprohormone. Preproinsulin is linearly composed of a signal peptide, as most secretory and membrane proteins, a B chain, a C peptide and an A chain (**Figure 2A**). Once the signal peptide is translated, it binds the signal recognition particle that docks to its receptor within the ER membrane and thus allows the direct translocation of newly synthesized insulin into the ER (**Figure 1C**). As the translation continues, the signal peptide is cleaved by signal peptidase (SP) and undergoes intramembrane proteolysis by the signal peptide peptidase (SPP) (47) while the nascent proinsulin is released in the ER. Proinsulin is further cleaved by endopeptidases that remove the C peptide, and folded by chaperones that favor the formation of two disulfide bonds between the B and the A chain, leading to mature insulin stored into granules until its exocytosis. The biosynthesis and secretion of insulin are two distinct and thoroughly regulated processes. Secretion is triggered by a higher glucose concentration (48, 49), meaning that insulin biosynthesis is continually active even at normal glucose concentrations and that most regulation occurs at the secretion stage. This feature has two consequences for APP. First, high numbers of free signal peptides are continually cleaved and released in the cytosol and ER. Second, this high bio-synthetic rate leads to a high number of misfolded insulin molecules, estimated to represent up to 20% of insulin production in normal, unstressed, beta cells (7).

The fate of post-cleavage signal peptides is poorly described. The most studied instance is that of non-classical HLA-E molecules, which present signal-peptide sequences derived from classical HLA-A, -B and -C molecules (50). Surface

peptide-loaded HLA-E complexes serve as a marker of proper MHC-I expression and prevent NK-cell activation and killing (51). Even though the presentation of HLA signal sequences by HLA-E is altered when TAP or tapasin are inhibited or knocked-out (52, 53), most reports suggest that the MHC-I presentation of signal peptides is a TAP- and proteasome-independent process (47, 54, 55). The presentation of signal-peptide-derived sequences is also increasingly recognized for classical MHC-I molecules, as described for tumor epitopes (56) and beta-cell peptides (23). We identified several HLA-A2-restricted sequences presented by beta cells and originating from the signal peptide of insulin granule proteins such as insulin, urocortin-3 (UCN3), islet amyloid polypeptide (IAPP) and proconvertase 2 (PCKS2), although the latter was not recognized by CD8+ T cells (23). Interestingly, preproinsulin epitopes restricted for HLA-A2 (23) and other MHC-I variants (57) map to the short (24 aa) signal peptide region. Focused investigations have therefore been conducted to characterize the generation of preproinsulin signal peptides for MHC-I presentation (57, 58). Using K562 cells transfected with a single MHC-I allele and insulin gene, different signal peptide products were found bound to each variant tested, and the fate of insulin signal peptide sequences was dependent on their ER intramembrane location. Indeed, after SPP cleavage within the ER membrane, N-terminal peptides are more likely to be released in the cytosol (**Figure 1C**). These cytosolic free peptides can follow the conventional APP route and be translocated to the ER through TAP transport and subsequently loaded on nascent MHC-I molecules after ERAP1 trimming. On the opposite, upon SPP cleavage, C-terminal peptides are directly released in the ER lumen and loaded on MHC-I molecules, in a TAP- and proteasome-independent manner (57, 58). It is unknown how a peptide can be loaded into MHC-I molecules without TAP and whether ERAP1 trimming is always required, as is the case for a PPI_{15-24} peptide (59). TAPBPR is the chaperone that likely replaces TAP, as tapasin and TAPBPR are mutually exclusive in their binding to MHC-I (55). One of the enzymes suspected to assist peptide trimming in the trans-Golgi network (TGN) is furin. Furin is a proconvertase that is known to traffic from the TGN to the cell surface *via* the vesicular pathway, where it can cleave peptides and proteins (60). Besides being involved in the maturation of secreted proteins such as those of the granule, furin has been shown to process antigens and direct them to the secretory route for MHC-I presentation in a TAP-independent manner (61). Similarly to another proconvertase from the same protein family (PC7) that has been shown to rescue unstable HLA-B51 complexes (62), furin could be crucial in post-ER stabilization of peptide-MHC-I binding. Indeed, resident TGN enzymes like furin could ensure an allele-dependent stability of these complexes against the destabilizing effect of the slightly acidic pH of the TGN. The mono-allelic HLA-I-expressing cells used in these studies may however introduce some bias and might not be representative of (multi-allelic) human beta cells. Consistent with their functional pH range, furin could also be involved in peptide exchanges occurring in the endosomes.

Collectively, the considerable translation rate of insulin and other granule proteins may provide a major source of aAPP pathways already within resting beta cells, through the release of high amounts of signal peptides and further protein processing.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

Beta cells are very active cells with a specialized secretory activity. On one hand, this activity requires unique recycling pathways to maintain homeostasis. On the other, it is a burden that makes beta cells more vulnerable to additional ER stress. These properties may enhance the contribution of MHC-I aAPP pathways in beta cells. Although direct evidence is missing, data from other cellular models make this possibility plausible and invite further investigations. Gaps in knowledge that need to be filled include: 1) the relative contribution of each APP pathway to the beta-cell immunopeptidome under basal and stressed conditions; 2) how these pathways are modulated in the pro-inflammatory environment of insulitis; and 3) whether the same antigenic peptides are generated by professional APCs, a required step to prime naïve T cells in draining lymph nodes before their homing to the pancreas.

AUTHOR CONTRIBUTIONS

AC and RM wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Work in our Laboratory is supported by The Leona M. and Harry B. Helmsley Charitable Trust (1901–03689), JDRF (2-SRA-2016-164-Q-R), the *Agence Nationale de la Recherche* (ANR-19-CE15-0014-01), the *Fondation pour la Recherche Medicale* (EQU20193007831), and the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, the European Federation of Pharmaceutical Industries and Associations, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust.

ACKNOWLEDGMENTS

We thank Dr Loredana Saveanu (Université de Paris, Centre de Recherche sur l'Inflammation, INSERM U1149, CNRS ERL8252) for her critical reviewing of the manuscript and expert advice.

REFERENCES

1. Pugliese A. Autoreactive T cells in type 1 diabetes. *J Clin Invest* (2017) 127:2881–91. doi: 10.1172/JCI94549
2. Mallone R, Eizirik DL. Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? *Diabetologia* (2020) 63:1999–2006. doi: 10.1007/s00125-020-05176-7
3. Battaglia M, Ahmed S, Anderson MS, Atkinson MA, Becker D, Bingley PJ, et al. Introducing the Endotype Concept to Address the Challenge of Disease Heterogeneity in Type 1 Diabetes. *Diabetes Care* (2020) 43:5–12. doi: 10.2337/dc19-0880
4. Leete P, Oram RA, McDonald TJ, Shields BM, Ziller CTIGI study team, et al. Studies of insulin and proinsulin in pancreas and serum support the existence of aetiopathological endotypes of type 1 diabetes associated with age at diagnosis. *Diabetologia* (2020) 63:1258–67. doi: 10.1007/s00125-020-05115-6
5. Carré A, Richardson SJ, Larger E, Mallone R. Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? *Diabetologia* (2021) 64:15–25. doi: 10.1007/s00125-020-05298-y
6. Blum JS, Wearsch PA, Cresswell P. Pathways of Antigen Processing. *Annu Rev Immunol* (2013) 31:443–73. doi: 10.1146/annurev-immunol-032712-095910
7. Sun J, Cui J, He Q, Chen Z, Arvan P, Liu M. Proinsulin misfolding and endoplasmic reticulum stress during the development and progression of diabetes. *Mol Aspects Med* (2015) 42:105–18. doi: 10.1016/j.mam.2015.01.001
8. Engin F, Yermalovich A, Nguyen T, Ngyuen T, Hummasti S, Fu W, et al. Restoration of the unfolded protein response in pancreatic β cells protects mice against type 1 diabetes. *Sci Transl Med* (2013) 5:211ra156. doi: 10.1126/scitranslmed.3006534
9. Eizirik DL, Pasquali L, Cnop M. Pancreatic β-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. *Nat Rev Endocrinol* (2020) 16:349–62. doi: 10.1038/s41574-020-0355-7
10. Eizirik DL, Miani M, Cardozo AK. Signalling danger: endoplasmic reticulum stress and the unfolded protein response in pancreatic islet inflammation. *Diabetologia* (2013) 56:234–41. doi: 10.1007/s00125-012-2762-3
11. Craig ME, Kim KW, Isaacs SR, Penno MA, Hamilton-Williams EE, Couper JJ, et al. Early-life factors contributing to type 1 diabetes. *Diabetologia* (2019) 62:1823–34. doi: 10.1007/s00125-019-4942-x
12. Dean PM. Ultrastructural morphometry of the pancreatic beta-cell. *Diabetologia* (1973) 9:115–9. doi: 10.1007/BF01230690
13. Muralidharan C, Conteh AM, Marasco MR, Crowder JJ, Kuipers J, de Boer P, et al. Pancreatic beta cell autophagy is impaired in type 1 diabetes. *Diabetologia* (2021) 64:865–77. doi: 10.1007/s00125-021-05387-6
14. Marasco MR, Conteh AM, Reissaus CA, Cupit JE, Appleman EM, Mirmira RG, et al. Interleukin-6 Reduces β-Cell Oxidative Stress by Linking Autophagy With the Antioxidant Response. *Diabetes* (2018) 67:1576–88. doi: 10.2337/db17-1280
15. Trivedi PC, Bartlett JJ, Pulinilkunnil T. Lysosomal Biology and Function: Modern View of Cellular Debris Bin. *Cells* (2020) 9:1131. doi: 10.3390/cells9051131
16. Soleimanpour SA, Gupta A, Bakay M, Ferrari AM, Groff DN, Fadista J, et al. The Diabetes Susceptibility Gene Clec16a Regulates Mitophagy. *Cell* (2014) 157:1577–90. doi: 10.1016/j.cell.2014.05.016
17. Hakonarson H, Grant SFA, Bradfield JP, Marchand L, Kim CE, Glessner JT, et al. A genome-wide association study identifies KIAA0350 as a type 1 diabetes gene. *Nature* (2007) 448:591–4. doi: 10.1038/nature06010
18. Grant BD, Donaldson JG. Pathways and mechanisms of endocytic recycling. *Nat Rev Mol Cell Biol* (2009) 10:597–608. doi: 10.1038/nrm2755
19. Montealegre S, van Endert PM. Endocytic Recycling of MHC Class I Molecules in Non-professional Antigen Presenting and Dendritic Cells. *Front Immunol* (2019) 9:3098:3098. doi: 10.3389/fimmu.2018.03098
20. Ilca FT, Neerincx A, Hermann C, Marcu A, Stevanović S, Deane JE, et al. TAPBPR mediates peptide dissociation from MHC class I using a leucine lever. *eLife* (2018) 7:e40126. doi: 10.7554/eLife.40126
21. Jahreiss L, Menzies FM, Rubinsztein DC. The Itinerary of Autophagosomes: From Peripheral Formation to Kiss-and-Run Fusion with Lysosomes. *Traffic* (2008) 9:574–87. doi: 10.1111/j.1600-0854.2008.00701.x
22. Yorimitsu T, Klionsky DJ. Eating the endoplasmic reticulum: quality control by autophagy. *Trends Cell Biol* (2007) 17:279–85. doi: 10.1016/j.tcb.2007.04.005
23. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze J-V, Nigi L, et al. Conventional and Neo-Antigenic Peptides Presented by β Cells Are Targeted by Circulating Naïve CD8+ T Cells in Type 1 Diabetic and Healthy Donors. *Cell Metab* (2018) 28:946–60. doi: 10.1016/j.cmet.2018.07.007
24. Bassani-Sternberg M, Pletscher-Frankild S, Jensen LJ, Mann M. Mass spectrometry of human leukocyte antigen class I peptidomes reveals strong effects of protein abundance and turnover on antigen presentation. *Mol Cell Proteomics* (2015) 14:658–73. doi: 10.1074/mcp.M114.042812
25. Azoury ME, Tarayrah M, Afonso G, Pais A, Colli ML, Maillard C, et al. Peptides Derived From Insulin Granule Proteins are Targeted by CD8+ T Cells Across MHC Class I Restrictions in Humans and NOD Mice. *Diabetes* (2020) 69:2678–90. doi: 10.2337/db20-0013
26. James EA, Mallone R, Kent SC, DiLorenzo TP. T-Cell Epitopes and Neo-epitopes in Type 1 Diabetes: A Comprehensive Update and Reappraisal. *Diabetes* (2020) 69:1311–35. doi: 10.2337/dbi19-0022
27. Wan X, Zinselmeyer BH, Zakharov PN, Vomund AN, Taniguchi R, Santambrogio L, et al. Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature* (2018) 560:107–11. doi: 10.1038/s41586-018-0341-6
28. Sims EK, Chaudhry Z, Watkins R, Syed F, Blum J, Ouyang F, et al. Elevations in the Fasting Serum Proinsulin–to–C-Peptide Ratio Precede the Onset of Type 1 Diabetes. *Diabetes Care* (2016) 39:1519–26. doi: 10.2337/dc15-2849
29. Rodriguez-Calvo T, Zapardiel-Gonzalo J, Amirian N, Castillo E, Lajevardi Y, Krogvold L, et al. Increase in Pancreatic Proinsulin and Preservation of β-Cell Mass in Autoantibody-Positive Donors Prior to Type 1 Diabetes Onset. *Diabetes* (2017) 66:1334–45. doi: 10.2337/db16-1343
30. Wasserfall C, Nick HS, Campbell-Thompson M, Beachy D, Haataja L, Kusmartseva I, et al. Persistence of Pancreatic Insulin mRNA Expression and Proinsulin Protein in Type 1 Diabetes Pancreata. *Cell Metab* (2017) 26:568–75. doi: 10.1016/j.cmet.2017.08.013
31. Sims EK, Bahnson HT, Nyalwidhe J, Haataja L, Davis AK, Speake C, et al. Proinsulin Secretion Is a Persistent Feature of Type 1 Diabetes. *Diabetes Care* (2019) 42:258–64. doi: 10.2337/dc17-2625
32. Sims EK, Syed F, Nyalwidhe J, Bahnson HT, Haataja L, Speake C, et al. Abnormalities in proinsulin processing in islets from individuals with longstanding T1D. *Transl Res* (2019) 213:90–9. doi: 10.1016/j.trsl.2019.08.001
33. Benkahla MA, Sabouri S, Kiosses WB, Rajendran S, Quesada-Masachs E, von Herrath M. HLA class I hyper-expression unmasks beta cells but not alpha cells to the immune system in pre-diabetes. *J Autoimmun* (2020) 119:102628. doi: 10.1016/j.jaut.2021.102628
34. Vieira E, Salehi A, Gylfe E. Glucose inhibits glucagon secretion by a direct effect on mouse pancreatic alpha cells. *Diabetologia* (2007) 50:370–9. doi: 10.1007/s00125-006-0511-1
35. Marroqui L, Masini M, Merino B, Grieco FA, Millard I, Dubois C, et al. Pancreatic α Cells are Resistant to Metabolic Stress-induced Apoptosis in Type 2 Diabetes. *EBioMedicine* (2015) 2:378–85. doi: 10.1016/j.ebiom.2015.03.012
36. Wyatt RC, Lanzoni G, Russell MA, Gerling I, Richardson SJ. What the HLA-I!—Classical and Non-classical HLA Class I and Their Potential Roles in Type 1 Diabetes. *Curr Diabetes Rep* (2019) 19:159. doi: 10.1007/s11892-019-1245-z
37. Himuro M, Miyatsuka T, Suzuki L, Miura M, Katahira T, Goto H, et al. Cellular Autophagy in α Cells Plays a Role in the Maintenance of Islet Architecture. *J Endocr Soc* (2019) 3:1979–92. doi: 10.1210/js.2019-00075
38. Sheng Q, Xiao X, Prasadan K, Chen C, Ming Y, Fusco J, et al. Autophagy protects pancreatic beta cell mass and function in the setting of a high-fat and high-glucose diet. *Sci Rep* (2017) 7:16348. doi: 10.1038/s41598-017-16485-0
39. Dalet A, Stroobant V, Vigneron N, Van den Eynde BJ. Differences in the production of spliced antigenic peptides by the standard proteasome and the immunoproteasome. *Eur J Immunol* (2011) 41:39–46. doi: 10.1002/eji.201040750
40. Lundh M, Bugliani M, Dahlby T, Chou DH-C, Wagner B, Ghiasi SM, et al. The immunoproteasome is induced by cytokines and regulates apoptosis in human islets. *J Endocrinol* (2017) 233:369–79. doi: 10.1530/JOE-17-0110
41. Seifert U, Bialy LP, Ebstein F, Bech-Otschir D, Voigt A, Schröter F, et al. Immunoproteasomes Preserve Protein Homeostasis upon Interferon-Induced Oxidative Stress. *Cell* (2010) 142:613–24. doi: 10.1016/j.cell.2010.07.036
42. Zaiss DMW, Bekker CPJ, Gröne A, Lie BA, Sijts AJAM. Proteasome immunosubunits protect against the development of CD8 T-cell-mediated

autoimmune diseases. *J Immunol* (2011) 187:2302–9. doi: 10.4049/jimmunol.1101003
43. Mishto M, Liepe J, Textoris-Taube K, Keller C, Henklein P, Weberruß M, et al. Proteasome isoforms exhibit only quantitative differences in cleavage and epitope generation. *Eur J Immunol* (2014) 44:3508–21. doi: 10.1002/eji.201444902
44. Winter MB, La Greca F, Arastu-Kapur S, Caiazza F, Cimermancic P, Buchholz TJ, et al. Immunoproteasome functions explained by divergence in cleavage specificity and regulation. *eLife* (2017) 6:e27364. doi: 10.7554/eLife.27364
45. Marroqui L, Dos Santos RS, Op de Beeck A, Coomans de Brachène A, Marselli L, Marchetti P, et al. Interferon-α mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes. *Diabetologia* (2017) 60:656–67. doi: 10.1007/s00125-016-4201-3
46. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, et al. Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Sci Immunol* (2018) 3:eaao4013. doi: 10.1126/sciimmunol.aao4013
47. Martoglio B, Dobberstein B. Signal sequences: more than just greasy peptides. *Trends Cell Biol* (1998) 8:410–5. doi: 10.1016/S0962-8924(98)01360-9
48. Alarcón C, Lincoln B, Rhodes CJ. The biosynthesis of the subtilisin-related proprotein convertase PC3, but no that of the PC2 convertase, is regulated by glucose in parallel to proinsulin biosynthesis in rat pancreatic islets. *J Biol Chem* (1993) 268:4276–80. doi: 10.1016/S0021-9258(18)53606-1
49. Schuit FC, In't Veld PA, Pipeleers DG. Glucose stimulates proinsulin biosynthesis by a dose-dependent recruitment of pancreatic beta cells. *Proc Natl Acad Sci USA* (1988) 85:3865–9. doi: 10.1073/pnas.85.11.3865
50. Braud V, Jones EY, McMichael A. The human major histocompatibility complex class Ib molecule HLA-E binds signal sequence-derived peptides with primary anchor residues at positions 2 and 9. *Eur J Immunol* (1997) 27:1164–9. doi: 10.1002/eji.1830270517
51. Braud VM, Allan DSJ, O'Callaghan CA, Söderström K, D'Andrea A, Ogg GS, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* (1998) 391:795–9. doi: 10.1038/35869
52. Lampen MH, Hassan C, Sluijter M, Geluk A, Dijkman K, Tjon JM, et al. Alternative peptide repertoire of HLA-E reveals a binding motif that is strikingly similar to HLA-A2. *Mol Immunol* (2013) 53:126–31. doi: 10.1016/j.molimm.2012.07.009
53. Celik AA, Kraemer T, Huyton T, Blasczyk R, Bade-Döding C. The diversity of the HLA-E-restricted peptide repertoire explains the immunological impact of the Arg107Gly mismatch. *Immunogenetics* (2016) 68:29–41. doi: 10.1007/s00251-015-0880-z
54. Wei ML, Cresswell P. HLA-A2 molecules in an antigen-processing mutant cell contain signal sequence-derived peptides. *Nature* (1992) 356:443–6. doi: 10.1038/356443a0
55. Oliveira CC, van Hall T. Alternative Antigen Processing for MHC Class I: Multiple Roads Lead to Rome. *Front Immunol* (2015) 6:298:298. doi: 10.3389/fimmu.2015.00298
56. Hage FE, Stroobant V, Vergnon I, Baurain J-F, Echchakir H, Lazar V, et al. Preprocalcitonin signal peptide generates a cytotoxic T lymphocyte-defined tumor epitope processed by a proteasome-independent pathway. *Proc Natl Acad Sci USA* (2008) 105:10119–24. doi: 10.1073/pnas.0802753105
57. Kronenberg-Versteeg D, Eichmann M, Russell MA, de Ru A, Hehn B, Yusuf N, et al. Molecular Pathways for Immune Recognition of Preproinsulin Signal Peptide in Type 1 Diabetes. *Diabetes* (2018) 67:687–96. doi: 10.2337/db17-0021
58. Skowera A, Ellis RJ, Varela-Calviño R, Arif S, Huang GC, Van-Krinks C, et al. CTLs are targeted to kill beta cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope. *J Clin Invest* (2008) 118:3390–402. doi: 10.1172/JCI35449
59. Thomaidou S, Kracht MJL, van der Slik A, Laban S, de Koning EJ, Carlotti F, et al. β-Cell Stress Shapes CTL Immune Recognition of Preproinsulin Signal Peptide by Posttranscriptional Regulation of Endoplasmic Reticulum Aminopeptidase 1. *Diabetes* (2020) 69:670–80. doi: 10.2337/db19-0984
60. Thomas G. Furin At The Cutting Edge: From Protein Traffic To Embryogenesis And Disease. *Nat Rev Mol Cell Biol* (2002) 3:753–66. doi: 10.1038/nrm934
61. Medina F, Ramos M, Iborra S, de León P, Rodríguez-Castro M, Del Val M. Furin-processed antigens targeted to the secretory route elicit functional TAP1-/-CD8+ T lymphocytes in vivo. *J Immunol* (2009) 183:4639–47. doi: 10.4049/jimmunol.0901356
62. Leonhardt RM, Fiegl D, Rufer E, Karger A, Bettin B, Knittler MR. Post-Endoplasmic Reticulum Rescue of Unstable MHC Class I Requires Proprotein Convertase PC7. *J Immunol* (2010) 184:2985–98. doi: 10.4049/jimmunol.0900308

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Carré and Mallone. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Annex 4: Original article showing that citrullinated GRP78 peptides can be recognized by circulating CD8+ T cells

CD8+ T Cells Variably Recognize Native Versus Citrullinated GRP78 Epitopes in Type 1 Diabetes

Marie Eliane Azoury,[1] Fatoumata Samassa,[1] Mijke Buitinga,[2] Laura Nigi,[3] Noemi Brusco,[3] Aïsha Callebaut,[2] Matthieu Giraud,[4] Magali Irla,[5] Ana Ines Lalanne,[1] Alexia Carré,[1] Georgia Afonso,[1] Zhicheng Zhou,[1] Barbara Brandao,[1] Maikel L. Colli,[6] Guido Sebastiani,[3] Francesco Dotta,[3] Maki Nakayama,[7] Decio L. Eizirik,[6,8] Sylvaine You,[1] Sheena Pinto,[9] Mark J. Mamula,[10] Yann Verdier,[11] Joelle Vinh,[11] Soren Buus,[12] Chantal Mathieu,[2] Lut Overbergh,[2] and Roberto Mallone[1,13]

Diabetes 2021;70:2879–2891 | https://doi.org/10.2337/db21-0259

In type 1 diabetes, autoimmune β-cell destruction may be favored by neoantigens harboring posttranslational modifications (PTMs) such as citrullination. We studied the recognition of native and citrullinated glucose-regulated protein (GRP)78 peptides by CD8+ T cells. Citrullination modulated T-cell recognition and, to a lesser extent, HLA-A2 binding. GRP78-reactive CD8+ T cells circulated at similar frequencies in healthy donors and donors with type 1 diabetes and preferentially recognized either native or citrullinated versions, without cross-reactivity. Rather, the preference for native GRP78 epitopes was associated with CD8+ T cells cross-reactive with bacterial mimotopes. In the pancreas, a dominant GRP78 peptide was instead preferentially recognized when citrullinated. To further clarify these recognition patterns, we considered the possibility of citrullination in the thymus. Citrullinating peptidylarginine deiminase (Padi) enzymes were expressed in murine and human medullary epithelial cells (mTECs), with citrullinated proteins detected in murine mTECs. However, Padi2 and Padi4 expression was diminished in mature mTECs from NOD mice versus C57BL/6 mice. We conclude that, on one hand, the CD8+ T cell preference for native GRP78 peptides may be shaped by cross-reactivity with bacterial mimotopes. On the other hand, PTMs may not invariably favor loss of tolerance because thymic citrullination, although impaired in NOD mice, may drive deletion of citrulline-reactive T cells.

Islet-reactive $CD8^+$ T cells are held as the final mediators of β-cell destruction (1). Surprisingly, they circulate at similar frequencies in most individuals, irrespective of type 1 diabetes status (2–4). This is true for known islet antigens and for novel ones, e.g., urocortin 3 (UCN3), secretogranin-5, and proconvertase-2 (3,4). The frequency of most of these islet epitope-reactive T cells falls in a relatively narrow range of ~1–50/10^6 $CD8^+$ T cells (0.0001–0.005%), which overlaps with that of naïve $CD8^+$ T cells recognizing viral peptides in virus seronegative donors (5,6). In agreement with this frequency overlap, circulating islet-reactive $CD8^+$ T cells are largely naïve in both healthy donors and

[1]Université de Paris, Institut Cochin, CNRS, INSERM, Paris, France
[2]Laboratory of Clinical and Experimental Endocrinology, KU Leuven, Leuven, Belgium
[3]Toscana Life Sciences, Diabetes Unit and Fondazione Umberto di Mario ONLUS, Department of Medicine, Surgery and Neuroscience, University of Siena, Siena, Italy
[4]Centre de Recherche en Transplantation et Immunologie, INSERM UMR1064, Université de Nantes, Nantes, France
[5]Centre d'Immunologie de Marseille-Luminy, INSERM, CNRS, Aix-Marseille University, Marseille, France
[6]Medical Faculty, Center for Diabetes Research and Welbio, Université Libre de Bruxelles, Brussels, Belgium
[7]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO
[8]Indiana Biosciences Research Institute, Indianapolis, IN
[9]Division of Developmental Immunology, Deutsches Krebsforschungszentrum, Heidelberg, Germany
[10]Yale University School of Medicine, New Haven, CT
[11]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS UMR8249, Paris, France
[12]Department of International Health, Immunology and Microbiology, Panum Institute, Copenhagen, Denmark
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, Paris, France

Corresponding author: Roberto Mallone, roberto.mallone@inserm.fr

Received 28 March 2021 and accepted 17 September 2021

This article contains supplementary material online at https://doi.org/10.2337/figshare.16654669.

M.E.A. and F.S. contributed equally to this work.

© 2021 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at https://www.diabetesjournals.org/content/license.

IMMUNOLOGY AND TRANSPLANTATION
Downloaded from http://diabetesjournals.org/diabetes/article-pdf/70/12/2879/634607/db210259.pdf by guest on 12 January 2022

donors with type 1 diabetes (2–4), suggesting a limited recruitment in the autoimmune process. The fraction engaged in this process seems instead sequestered in the pancreas, where the same T cells are enriched in patients with type 1 diabetes and display a $CD45RO^+$ phenotype, compatible with a prior antigen encounter (2,3). Thus, a universal state of "benign" islet autoimmunity exists for circulating $CD8^+$ T cells, to a much larger extent than previously appreciated (1,7). In line with recent reports (6), we demonstrated that this benign autoimmunity is imprinted in the thymus due to a marginal deletion of autoreactive $CD8^+$ T cells, since the presence or absence of ectopic islet antigen presentation in thymic medullary epithelial cells (mTECs) does not significantly modulate this process (2,3).

It is unknown whether this novel paradigm of benign autoimmunity applies to β-cell antigens undergoing posttranslational modifications (PTMs). PTMs occurring in peripheral tissues, especially under inflammatory conditions, may favor priming of T cell precursors that have escaped thymic deletion, due to their selection against native peptides with different physicochemical properties (8). One posttranslationally modified antigen is glucose-regulated protein (GRP)78, which is citrullinated on arginine (R) residues in inflamed mouse and human islets (9,10) by peptidyl arginine deiminase enzymes (PADI; Padi in the mouse). GRP78 (also known as BiP and HSPA5) is an endoplasmic reticulum (ER) chaperone of the heat shock protein 70 family. It is involved in anabolic pathways, i.e., protein (re)folding and translocation of secretory proteins, and catabolic pathways, i.e., the import of polypeptides—especially hydrophobic ones—inside the ER, the ER-associated degradation pathway, peptide export outside the ER for proteasome degradation, and the regulation of the unfolded protein response (11). Citrullinated GRP78 is translocated to the plasma membrane and secreted, triggering β-cell apoptosis (9,12).

We first reported that the murine ortholog of GRP78 (Grp78) is targeted by T cells and autoantibodies (aAb) in NOD mice (9). While interferon-γ (IFN-γ) T cell responses were only observed against citrullinated Grp78, aAb were detected against both native and citrullinated Grp78, but the latter was preferentially recognized, with higher aAb titers in NOD mice than in nonautoimmune strains (9). Translating to humans, a $CD4^+$ T-cell line grown from type 1 diabetic islets responded to a $GRP78_{292–305}$ peptide only when carrying an arginine-to-citrulline (R→X) modification (13). aAb against both native GRP78 (R-GRP78) and citrullinated (X-GRP78) GRP78 were detected at increased titers in individuals with type 1 diabetes versus age-matched healthy individuals, but again X-GRP78 was preferentially recognized (10). $CD4^+$ T cells reactive to an HLA-DR4 binding $X\text{-}GRP78_{498–512}$ peptide displayed higher frequencies in donors with type 1 diabetes versus healthy donors, but these frequencies were similar to those of $CD4^+$ T cells recognizing $R\text{-}GRP78_{498–512}$. Importantly, this $GRP78_{498–512}$ epitope features an R510 residue that is citrullinated in human islets exposed in vitro to inflammatory cytokines (10). Collectively, these results suggest that $CD4^+$ T cells recognize preferentially, but not exclusively, X-GRP78 over its native version.

We therefore studied the recognition of R- and X-GRP78 peptides restricted for HLA-A*02:01 (HLA-A2) by the $CD8^+$ T cells of donors with type 1 diabetes and healthy donors. Our results show that the preferential recognition of X-GRP78 versions is not the rule and that this recognition may be shaped by T cell cross-reactivity with bacterial mimotopes and by PADI-mediated citrullination in the thymus and islets.

RESEARCH DESIGN AND METHODS

HLA-A2 Binding Measurements

Experimental binding of peptides (>90% purity; SynPeptide) to HLA-A2 was measured by flow cytometry using biotin-tagged HLA-A2 monomers (immunAware) assembled as previously described (14), captured on streptavidin-coated beads (Spherotech), and revealed by an anti-β2m monoclonal antibody (Research Resource Identifier [RRID] AB_626748) and an Alexa Fluor (AF)488-conjugated goat anti-mouse IgG (RRID AB_2728715).

Donors and Specimens

Human studies were authorized under ethics approval no. S52697 (UZ Leuven), and written informed consent was obtained from all participants. Heparinized blood was drawn from $HLA\text{-}A*02:01^+$ donors: 16 aAb^+, insulin-treated adults with type 1 diabetes (median age 27 years [range 19–47]; 9 female and 7 male; median disease duration 4.3 years [1.7–6.0]); 11 healthy donors (median age 37 years [26–60]; 7 female and 4 male); and 8 adolescents with type 1 diabetes (median age 16 years [12–17]; 3 female and 5 male; median disease duration 4.2 years [1.0–12.0]). Ficoll-purified peripheral blood mononuclear cells (PBMCs) were stored in liquid nitrogen until use. Pancreatic tissue sections were provided by the Network for Pancreatic Organ donors with Diabetes (nPOD).

HLA-A2 Multimer Staining

Our combinatorial HLA-A2 multimer (MMr) assays, previously described (3,4), allow us to multiplex the detection of 15 different T-cell epitope specificities in the same PBMC sample. It is based on the use of unique fluorochrome pairs to label each peptide-loaded MMr, with subsequent detection of double-labeled MMr^+CD8^+ T cells with increased specificity. This format was modified so that each R-GRP78 MMr shared one fluorochrome with its corresponding X-GRP78 MMr, thus allowing selective gating of $CD8^+$ T cells recognizing only the R or X version (i.e., double MMr^+) or both (i.e., triple MMr^+). For example, the GRP78-261R and GRP78-261X peptides were loaded on phycoerythrin (PE)/BV786- and PE/PE-CF594–labeled MMr pairs, respectively, thus sharing the PE fluorochrome. GRP78-261R, -261X, and -261R/X MMr^+CD8^+ T cells

were, hence, defined as $PE^+/BV786^+$, $PE^+/PECF594^+$, and $PE^+/BV786^+/PECF594^+$, respectively. MMr staining and acquisition were performed as previously described (3). HLA-A2 binding epitope candidates without appreciable MMr staining provided negative controls. Data were analyzed with FlowJo software as detailed in Supplementary Figs. 1 and 2. In situ HLA-A2 MMr immunohistochemistry staining was performed as previously described (3).

mTEC Isolation

For quantitative RT-PCR (RT-qPCR) and proteomics, thymi from 6- to 8-week-old C57BL/6 and NOD mice were minced; gently agitated for release of excess thymocytes; digested with Liberase; filtered; panned on anti-CD90.2–coated plates; stained with Fixable Viability Dye eFluor 780, CD45-PE/Cy7, MHC class II (MHC-II)–AF488, EPCAM-BV421, BP-1–APC, *Ulex europaeus* agglutinin-1 (UEA-1–biotin), and streptavidin-PE; and sorted on a BD Influx instrument.

For RNA sequencing (RNAseq), thymi were digested with collagenase D and collagenase/dispase; depleted with CD45 MicroBeads; stained with CD45-PerCPCy5.5, Ly51-PE, and MHC-II–APC; and sorted on a BD FACSAria III instrument. These procedures are described in Supplementary Material.

RT-qPCR, Proteomics, and RNAseq

Mouse studies were performed under authorization no. P076/2020 (UZ Leuven). Islets were isolated from 6- and 10-week-old C57BL/6 and NOD mice as previously described (15). Total RNA was extracted from sorted mTECs or islets, with the Single Cell RNA Purification Kit (Norgen Biotek) and cDNA synthesized with SuperScript VILO (Invitrogen). RT-qPCR was performed on a StepOnePlus Real-Time PCR System (Applied Biosystems). The relative fold gene expression was calculated using the $\Delta\Delta$Ct method.

Liquid chromatography–tandem mass spectrometry and manual verification of citrullinated residues were performed essentially as previously described (16). RNAseq procedures have previously been published (17). These procedures are detailed in Supplementary Material.

Statistics

Significance was assessed with two-tailed tests with a cutoff value of $\alpha = 0.05$, as detailed for each figure and in Supplementary Material.

Data and Resource Availability

The mTEC RNAseq data sets have been deposited in the Gene Expression Omnibus (GEO) under accession nos. GSE140815, GSE140683, and GSE178456. All raw data are available from the corresponding author upon reasonable request. The study does not involve any noncommercial reagent or tool.

RESULTS

Citrullination of GRP78 Peptides Does Not Decrease HLA-A2 Binding Affinity

We selected six candidate epitopes by screening the whole GRP78 sequence for predicted HLA-A2 binders carrying R residues (Fig. 1*A*): $GRP78_{9–18}$ (GRP78-9 from hereon), $GRP78_{261–269}$ (GRP78-261), $GRP78_{298–307}$ (GRP78-298), $GRP78_{360–368}$ (GRP78-360), $GRP78_{435–443}$ (GRP78-435), and $GRP78_{502–511}$ (GRP78-502).

Since binding affinities were predicted for the native peptide sequence without citrullination, we experimentally verified binding using monomeric HLA-A2 molecules and a flow cytometry assay detecting peptide/HLA-A2/β2m complexes (4) (Fig. 1*B*). First, this assay verified the proper folding of the peptide/HLA-A2 monomers subsequently used for MMr synthesis. Second, it revealed strong (GRP78-261, -435, -298), moderate (GRP78-360, -9), and weak (GRP78-502) binders (Fig. 1*C* and *D*). The HLA-A2 binding of the citrullinated (X) versus native (R) versions was similar (GRP78-261, -435), marginally increased (GRP78-298, -502), or significantly increased (GRP78-360, -9) but never decreased. The increase in HLA-A2 binding was observed for those peptides in which the R→X substitution was introduced at the COOH-terminal P8-P9 HLA anchor positions.

Collectively, these data show that citrullination does not negatively affect HLA-A2 binding for the candidate epitopes identified.

Blood $CD8^+$ T Cells Preferentially Recognize Either Native or Citrullinated Peptide Versions and Display Similar Frequencies in Donors With Type 1 Diabetes and Healthy Donors

We subsequently set up combinatorial HLA-A2 MMr assays (3,4), with two purposes: first, to define the ex vivo frequency and naïve/antigen-experienced phenotype of GRP78-reactive $CD8^+$ T cells in donors with type 1 diabetes and healthy donors, and second, to investigate whether the native and citrullinated versions of any given peptide are recognized by the same T cell clonotypes. For this second purpose, we modified our previous assays so that the MMr pairs loaded with each R version shared one fluorochrome with the MMr pairs loaded with its corresponding X version (Supplementary Table 1), thus allowing selectively gating of $CD8^+$ T cells recognizing only the R or X version or both (Fig. 2*A* and *B*). A representative staining is shown in Supplementary Fig. 1, and the analysis strategy is detailed in Supplementary Fig. 2.

The results from adult donors are summarized in Fig. 2*C*. Recognition of the R and X versions was rather exclusive, with very few donors recognizing both. The native R version was preferentially recognized for peptides GRP78-9, -298, -360, and -435, while GRP78-261 and -502 were preferred in their X version. The R version was also preferred for those peptides whose citrullination at the COOH terminus enhanced HLA-A2 binding (GRP78-9, -360, and, to a lesser extent, -298), thus excluding that

A

*MKLSLVAA**MLLLLSAARA***AEEEDKKEDVGTVVGIDLGTTYSCVGVFKNGRVEIIAND
QGNRITPSYVAFTPEGERLIGDAAKNQLTSNPENTVFDAKRLIGRTWNDPSVQQDI
KFLPFKVVEKKTKPYIQVDIGGGQTKTFAPEEISAMVLTKMKETAEAYLGKKVTHAV
VTVPAYFNDAQRQATKDAGTIAGLNVMRIINEPTAAAIAYGLDKREGEKNILVFDLG
GGTFDVSLLTIDNGVFEVVATNGDTHLGGEDFDQ**RVMEHFIKL**YKKKTGKDVRKD
NRAVQKLRREVEKAKR**ALSSQHQARI**EIESFYEGEDFSETLTRAKFEELNMDLFRS
TMKPVQKVLEDSDLKKSDIDEI**VLVGGSTRI**PKIQQLVKEFFNGKEPSRGINPDEAV
AYGAAVQAGVLSGDQDTGDLVLLDVCPLTLGIETVGGVMT**KLIPRNTVV**PTKKSQIF
STASDNQPTVTIKVYEGERPLTKDNHLLGTFDLTGIPPAPRGVPQIEVTF**EIDVNGIL
RV**TAEDKGTGNKNKITITNDQNRLTPEEIERMVNDAEKFAEEDKKLKERIDTRNELE
SYAYSLKNQIGDKEKLGGKLSSEDKETMEKAVEEKIEWLESHQDADIEDFKAKKKE
LEEIVQPIISKLYGSAGPPPTGEEDTAEKDEL*

GRP78 peptide	Sequence	Predicted HLA-A2 binding		Experimental HLA-A2 binding (AU)	
		NetMHC (nM)	SYFPETHI	R isoform	X isoform
9-18	MLLLLSAARA	195	18	3	14
261-269	RVMEHFIKL	39	19	33	32
298-307	ALSSQHQARI	4483	22	12	24
360-368	VLVGGSTRI	300	23	2	7
435-443	KLIPRNTVV	38	25	25	29
502-511	EIDVNGILRV	9080	20	1	2

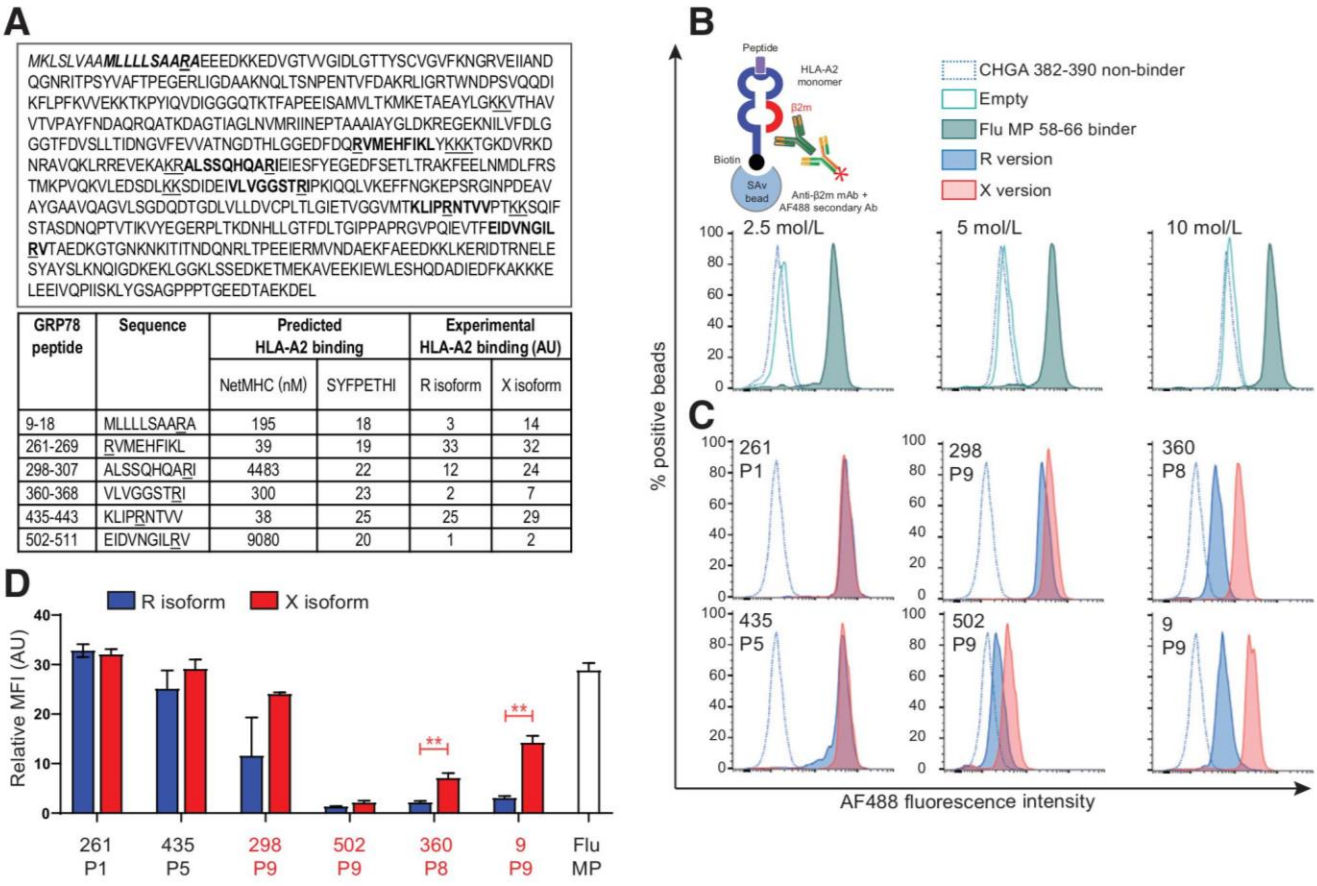


Figure 1—HLA-A2 binding of native (R) and citrullinated (X) GRP78 epitopes. *A*: Epitope candidates (shown in bold within the GRP78 sequence, UniProt identifier P11021) were selected in silico based on a NetMHC 4.0–predicted K_D ≤300 nmol/L or a SYFPETHI score ≥20 (www.cbs.dtu.dk/services/NetMHC and www.syfpeithi.de, respectively). For length variants of the same peptide, the one with the best binding affinity was retained. R residues comprised within the selected candidates and dibasic cleavage motifs are underlined. The GRP78 signal peptide is indicated in italics. The experimental HLA-A2 binding affinity measured for the R and X versions in the experiments shown in subsequent panels is reported in arbitrary units (AU). *B*: Schematic view and setup of the assay. Biotinylated monomeric HLA-A2 molecules are incubated with peptide and β2m. After capturing on streptavidin-coated beads, an anti-β2m antibody is added, followed by an AF488-labeled secondary antibody. The bead-associated fluorescence is therefore detected only if the test peptide supports the folding of the HLA-A2/β2m complex. Histograms (gated on single AF488$^+$ beads) depict representative flow cytometry staining obtained with increasing concentrations of HLA-A2 molecules folded with the nonbinding $CHGA_{382–390}$ (HPVGEADYF) and the binding Flu $MP_{58–66}$ (GILGFVFTL) peptides (negative and positive control, respectively). The 2.5 nmol/L concentration was retained for subsequent experiments based on the best separation of the positive and negative peaks. *C*: Representative staining of the tested GRP78 peptides in their native (blue) or citrullinated versions (red). The R/X position within the peptide sequence, either outside (P1, P5) or inside (P8, P9) the COOH-terminal region, is indicated in each histogram. *D*: Relative median fluorescence intensity (MFI) AU values for peptide–HLA-A2 complexes at 2.5 nmol/L, normalized to the $ChgA_{382–390}$ negative control peptide. Data are expressed as median ± SD of triplicate measurements, and the R/X position is indicated for each peptide. Ab, antibody; mAb, monoclonal antibody. $^{**}P \leq 0.003$ by Student *t* test.

this preference was simply driven by HLA-A2 binding affinity. Despite some differences in frequency, the preferred and exclusive recognition of the R or X version according to individual peptides was similar in adolescents with type 1 diabetes (Fig. 3).

As for other islet epitopes (2–4,18–20), the frequency of GRP78-reactive CD8$^+$ T cells was not higher in adults with type 1 diabetes than in healthy adults (Fig. 2*C*), with one exception noted for GRP78-435R. All positive GRP78-reactive CD8$^+$ T cell fractions displayed frequencies similar to those observed for the control preproinsulin $(PPI)_{6–14}$ islet epitope, confirming the reported range of ~1–50/10^6 CD8$^+$ T cells (2–4).

Collectively, these results show that, depending on the individual peptide, CD8$^+$ T cells exclusively recognize either the native or citrullinated version, with no significant cross-reactivity. Moreover, they circulate at similar frequencies in adults with type 1 diabetes and healthy adults.

GRP78 Peptides Preferentially Recognized in Their Native Version Are Cross-reactive With Gut Bacterial Mimotopes

Another recurrent feature of circulating CD8$^+$ T cells recognizing other islet epitopes is their largely naïve phenotype (2–4). GRP78-reactive CD8$^+$ T cells were also largely (>80%) naïve in most donors, but substantial antigen-

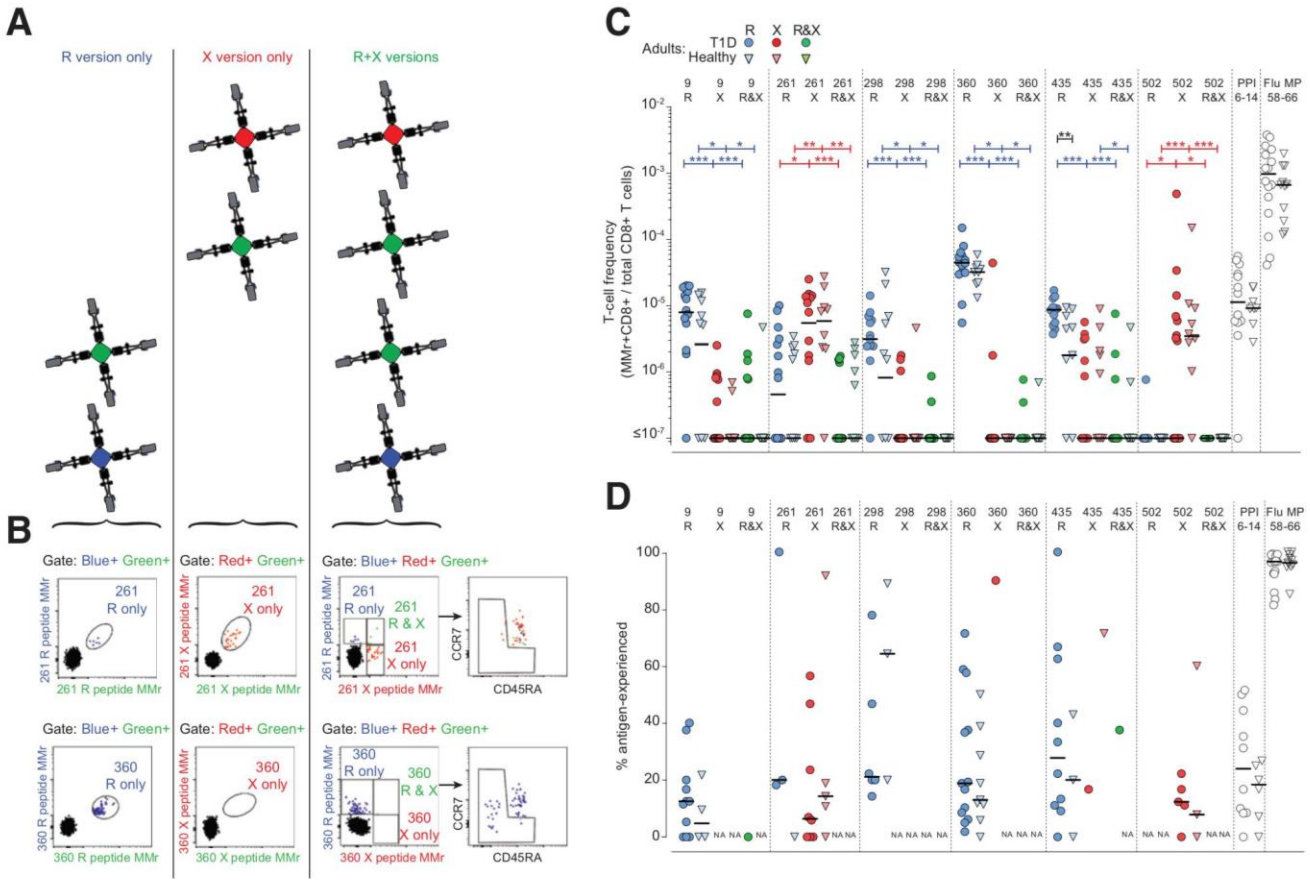


Figure 2—Blood CD8$^+$ T cells recognize either R- or X-GRP78 peptide versions and display similar frequencies in adults with type 1 diabetes and healthy adults. *A*: Schematic view of the combinatorial analysis applied to GRP78 MMr$^+$CD8$^+$ T cells. Each peptide is loaded onto HLA-A2 MMrs coupled with two different fluorochromes. Six fluorochromes are used to obtain 15 unique pairs. Consequently, each peptide-reactive cell population is stained with a unique pair of fluorochromes. Native (R) and citrullinated (X) versions of the same GRP78 peptide are stained with fluorochrome pairs that share one fluorochrome (green). With gating on double-MMr$^+$ events and exclusion of all other fluorochromes, cells recognizing only the R version or X version are visualized. With gating on triple-MMr$^+$ events, cells recognizing the R only, X only, or both versions are visualized. Subsequent staining with CD45RA and CCR7 allows characterization of the naïve/memory phenotype of each population. *B*: Representative examples for the GRP78-261 (top) and -360 (bottom) peptide. This strategy is detailed in Supplementary Figs. 1 and 2. *C*: Frequencies of MMr$^+$CD8$^+$ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X (red symbols) peptides, or to both (green symbols), in HLA-A2$^+$ adults with type 1 diabetes (T1D) (circles) (n = 16) and healthy adults (triangles) (n = 11). $PPI_{6–14}$ and Flu $MP_{58–66}$ peptides were included as controls. At least 0.13 × 10^6 CD8$^+$ T cells were counted for each donor (median 5.4 × 10^6 [range 0.13–280 × 10^6]). *D*: Percentage of antigen-experienced cells out of total MMr$^+$ cells for the GRP78 peptides depicted in *C*. Data points with <5 MMr$^+$ cells were excluded (median 13 MMr$^+$ cells [range 5–524] for GRP78 peptides). Bars show median values. NA, not available (i.e., <5 MMr$^+$ cells counted). *P < 0.05, **P < 0.01, ***P < 0.001 by Mann-Whitney U test.

experienced fractions were observed in some. This was noteworthy for epitopes GRP78-298R, -360R, and -435R preferentially recognized in their native version, which displayed >30% antigen-experienced fractions in >30% of adult donors (Fig. 2*D*). We therefore hypothesized that these R-GRP78–reactive CD8$^+$ T cells may cross-react with foreign antigens. A sequence homology search (Table 1) retrieved homologous bacterial sequences (mimotopes) for these three epitopes and for GRP78-502, which was instead predominantly recognized as citrullinated. Bacterial sequences were identical for GRP78-435R and, notably, GRP78-360R, which displayed particularly high cognate CD8$^+$ T-cell frequencies (Fig. 2*C*). These mimotopes mapped to proteins expressed by gut commensal bacteria (*Enterobacteriaceae*, *Proteobacteria*, *Clostridia*, and *Escherichia coli*), including the well-known DnaK chaperone (21). While the GRP78-360R– and GRP78-435R–identical bacterial sequences ("homotopes") are by definition cross-reactive, we investigated whether this was also the case for the GRP78-298R and -502X mimotopes. To this end, we used a combinatorial MMr strategy similar to the one applied to define R- and X-GRP78 cross-reactivity (Fig. 4*A*). Each of three PBMC aliquots was stained with two MMr pairs, and each MMr pair was loaded with the GRP78, mimotope, or control GRP78-360R or Flu $MP_{58–66}$ peptide. The two MMr pairs in each tube shared one fluorochrome with each other for visualization of cross-reactive T cells as triple-labeled MMr$^+$ events. A sizable fraction (24 ± 4% in healthy donors, 25 ± 14% in donors with type 1 diabetes) of CD8$^+$ T cells recognizing the GRP78-298R epitope

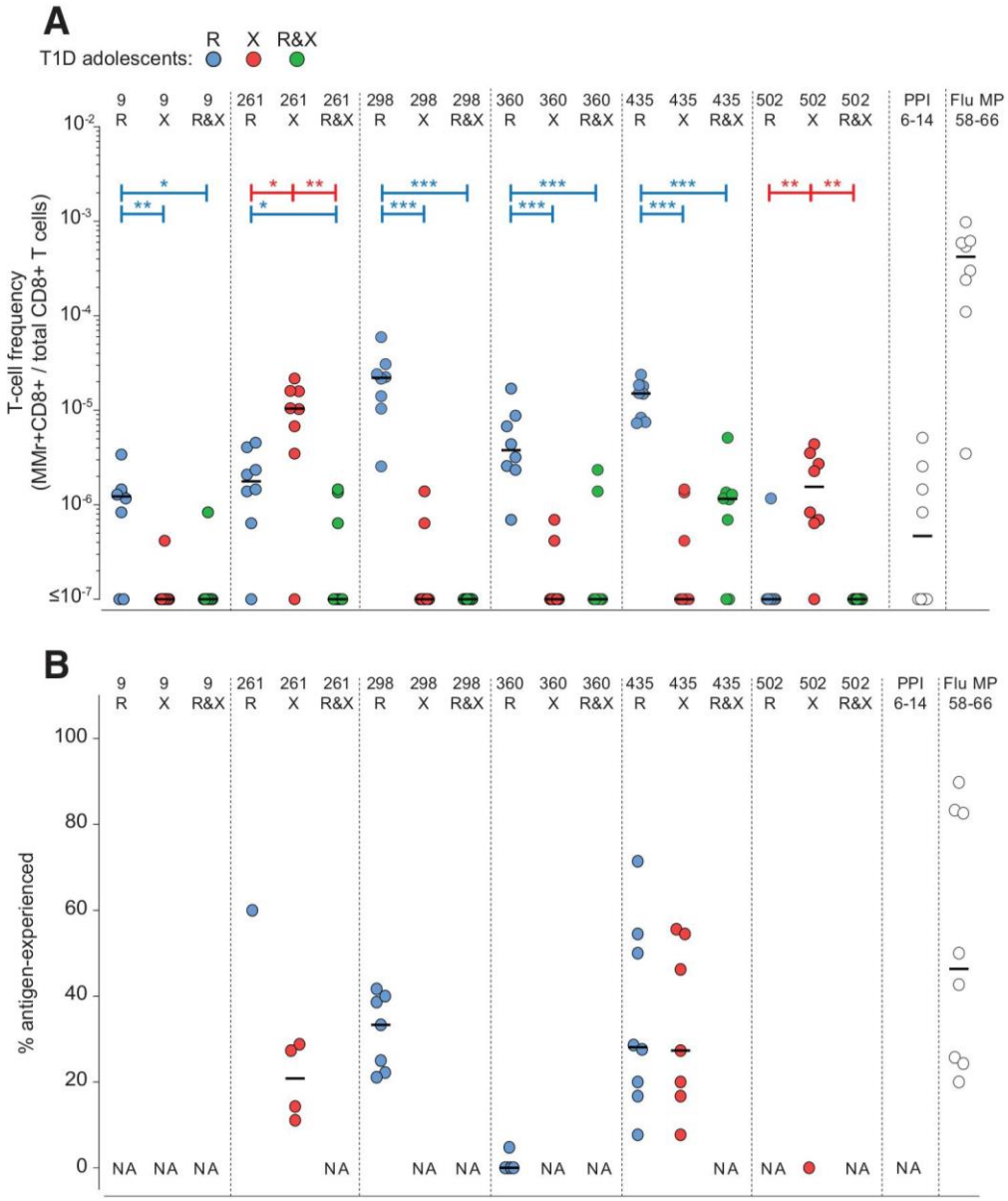


Figure 3—Blood CD8$^+$ T cells recognize either R- or X-GRP78 peptide versions in adolescents with type 1 diabetes. *A*: Frequencies of MMr$^+$CD8$^+$ cells reactive exclusively to the GRP78-R (blue symbols) or GRP78-X (red symbols) peptides, or to both (green symbols), in HLA-A2$^+$ adolescents with type 1 diabetes (T1D) (n = 8). $PPI_{6–14}$ and Flu $MP_{58–66}$ peptides were included as controls. At least 0.39 × 10^6 CD8$^+$ T cells were counted for each donor (median 1.1 × 10^6 [range 0.39–2.4 × 10^6]). *B*: Percentage of antigen-experienced cells out of total MMr$^+$ cells for the GRP78 peptides depicted in *A*. Data points with <5 MMr$^+$ cells were excluded (median 15 MMr$^+$ cells [range 5–115] for GRP78 peptides). Bars show median values. NA, not available (i.e., <5 MMr$^+$ cells counted). *P < 0.05, **P < 0.01, ***P < 0.001 by Mann-Whitney U test.

was cross-reactive with its bacterial mimotope (Fig. 4*B*, left), while this was not the case in analyses of cross-reactivity with the control GRP78-360R or Flu epitope (Fig. 4*B*, middle). As a positive control, GRP78-360R/GRP78-360R or Flu/Flu triple-MMr$^+$ events were correctly visualized as largely cross-reactive (Fig. 4*B*, right). This cross-reactivity was instead lower for CD8$^+$ T cells recognizing the GRP78-502X epitope (4 ± 3% of total in healthy donors, 13 ± 4% in donors with type 1 diabetes) (Fig. 4*C*).

Collectively, these results suggest that the preferential CD8$^+$ T cell recognition of some R-GRP78 versions is influenced by cross-reactivity with bacterial mimotopes.

Pancreas-Infiltrating CD8$^+$ T Cells Preferentially Recognize the Citrullinated GRP78-9X Epitope

While the preferential recognition of GRP78 native versions in peripheral blood may be favored by bacterial cross-reactivity, the situation may be different in the pancreas, where citrullination and its catalyzing enzymes

Table 1—GRP78 mimotopes from human intestinal commensal bacteria

GRP78-298R (R version preferred)	**ALSSQHQARI**
Enterobacteriaceae bacterium acetyl-CoA C-acyltransferase	*ALSSQH–ARI*
UniProt A0A3C2BKS9	ALSSQHKARI
GRP78-360R (R version preferred)	**VLVGGSTRI**
Proteobacteria DnaK chaperone	*VLVGGSTRI*
UniProt A0A2G6J6V3	VLVGGSTRI
GRP78-435R (R version preferred)	**KLIPRNTVV**
Clostridium species restriction endonuclease subunit S	*KLIPRNTVV*
UniProt A0A3R6J648, R7CB01, A0A373M041	KLIPRNTVV
GRP78-502X (X version preferred)	**EIDVNGILXV**
E. coli DnaK chaperone	*+ID+–GIL–V*
UniProt P0A6Y8	DIDADGILHV

Homologous bacterial sequences were searched with Basic Local Alignment Search Tool (BLAST). The alignment with GRP78 epitope sequences (in boldface type) is shown in italics, where – and + indicate nonconservative and conservative substitutions, respectively. The human GRP78 sequence is from UniProt P11021.

PADIs are upregulated by islet inflammation (9,10). We therefore performed in situ MMr staining on pancreas sections from nPOD donors (Supplementary Table 2) using the prototype epitope GRP78-9R/X, whose R version was preferentially recognized in peripheral blood. A representative GRP78-9X MMr staining on pancreas sections is shown in Fig. 5*A*. As reported for other islet antigen specificities (2,3,22), scattered GRP78-9X MMr^+ cells were detected in the islet vicinity and in the exocrine tissue. In donors with type 1 diabetes, they displayed significantly higher densities than control MMr^+ cells recognizing the melanocyte epitope $MelanA_{26–35}$ (Fig. 5*B*). However, GRP78-9X MMr^+ cells were not enriched in patients with type 1 diabetes compared with donors without diabetes, suggesting no disease specificity. The density of these GRP78-9X MMr^+ cells was lower than that of $ZnT8_{186–194}$ MMr^+ cells and similar to that of the $UCN3_{1–9}$ MMr^+ cells. GRP78-9X MMr^+ cells were significantly more abundant than their GRP78-9R MMr^+ counterparts in pooling of donors with type 1 diabetes and aAb^+ donors but not in pooling of control donors without diabetes and donors with type 2 diabetes. They were instead absent in duodenal mucosa sections from those with type 1 diabetes (Fig. 5*C*), despite the presence of $CD8^+$ T cells (not shown), thus confirming pancreas specificity. Analysis of a more limited set of samples for the epitope GRP78-360R/X, which, like GRP78-9R/X, was preferentially recognized in its R version in peripheral blood, also detected recognition of the X version in the pancreas of one donor with type 1 diabetes (Fig. 5*D*).

Collectively, these data indicate that citrullinated GRP78-9X rather than its native GRP78-9R version is recognized by $CD8^+$ T cells in the pancreas and that this pancreas-specific reactivity is not enriched in those with type 1 diabetes.

The Citrullinating Padi2 Enzyme Is Expressed in Murine mTECs Under the Control of Aire and Is Diminished in NOD Mice

To summarize, we observed that blood $CD8^+$ T cells preferentially recognize the native R-GRP78 versions for several epitopes, including GRP78-9R, while the citrullinated GRP78-9X version is preferentially recognized in the pancreas. Besides the potential effect of bacterial cross-reactivity, we asked how to reconcile these findings with the expected impact of central and peripheral tolerance mechanisms in shaping these responses. It is currently held that PTM recognition should be favored because PTMs may take place under inflammatory conditions in the periphery and not in the thymus (8). Hence, T cells recognizing citrullinated epitopes may be more likely to escape thymic deletion and may then be preferentially primed in the inflammatory milieu of insulitis that promotes citrullination.

To address this possibility, we compared the gene expression of *Padi2*, which is the predominant islet isoform (23), in the islets and thymus between NOD and nonautoimmune C57BL/6 mice. As previously reported (9), islet *Padi2* expression was negligible in C57BL/6 mice and higher in NOD mice (Fig. 6*A*). This expression increased with age and correlated with a difference in the Padi enzymatic activity (Fig. 6*B*). In mTECs (Fig. 6*C* and *D*; gating strategy in Supplementary Fig. 3*A* and *B*), *Padi2* and *Padi4* were substantially upregulated with maturation in C57BL/6 but not NOD mice and lower in NOD versus C57BL/6 mice for the mature MHC-II^{hi} mTECs driving T cell deletion. Mature MHC-II^{hi} mTECs were also less represented in NOD mice (Supplementary Fig. 3*C*). The Padi enzymatic activity from whole thymic tissue was not significantly different between the two strains (Fig. 6*B*), possibly reflecting the mixture of mTECs (both MHC-II^{lo} and MHC-II^{hi}) and other cells in these tissue preparations. Proteomics analysis of mTECs revealed citrullinated R residues in three proteins (Supplementary Fig. 4) (3 of 920 unique proteins detected [0.33%]). *PADI* gene expression was also observed in human mTECs (Supplementary Fig. 5) and was dominated by the *PADI2* isoform.

We next asked whether mTECs express *Padi* isoforms under the control of the autoimmune regulator (Aire) transcription factor. Expression of all *Padi* isoforms was

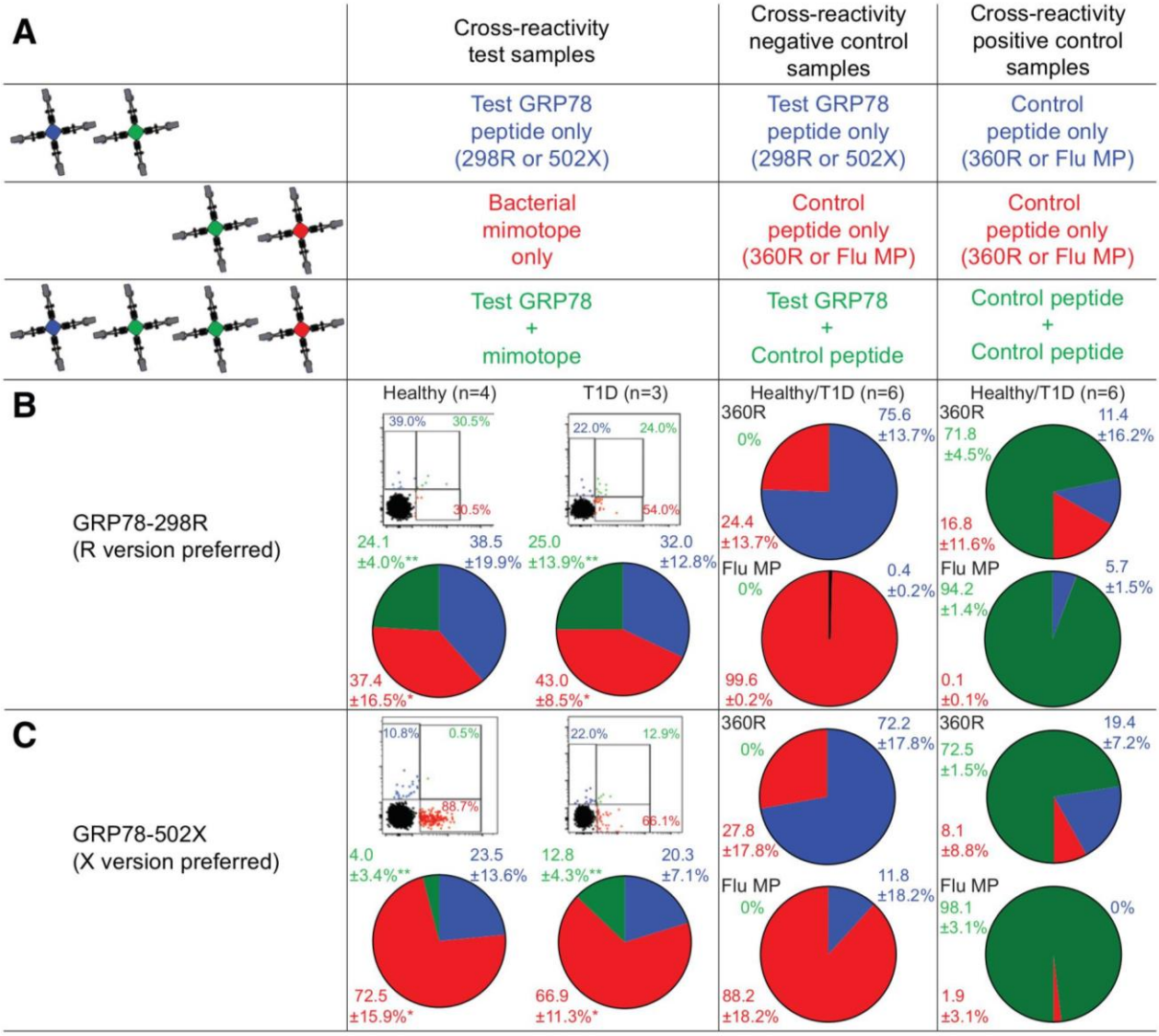


Figure 4—GRP78 peptides preferentially recognized in their native R version are cross-reactive with gut bacterial mimotopes. *A*: Schematic view of the combinatorial analysis applied. With a strategy similar to the one described in Fig. 2*A*, three PBMC aliquots were stained with two MMr pairs sharing one fluorochrome (green). With gating on double or triple MMr^+ events, cells recognizing only one peptide or both are visualized, respectively. *B* and *C*: Representative dot plots and cumulative results from healthy donors and donors with type 1 diabetes (pie charts) obtained for GRP78-298R (*B*), GRP78-502X (*C*), and their respective peptide mimotopes. The first column shows the results obtained on the first T cell aliquot (test sample) by crossing MMrs loaded with GRP78 and bacterial mimotope peptides, with the T cell fraction recognizing both highlighted in green. The second column displays the cross-reactivity between GRP78 peptide and GRP78-360R or Flu $MP_{58–66}$ epitope (negative control; second PBMC aliquot). The third column displays the cross-reactivity between GRP78-360R or Flu $MP_{58–66}$ peptides charged on both MMr pairs (positive control; third PBMC aliquot). Dot plots and pie charts display the percentage of MMr^+ cells binding either MMr (blue or red) or both (green). In dot plots, the MMr^- population is displayed in black for visualization of the position of MMr^+ events relative to it. $^{**}P = 0.016$ and $^{*}P = 0.031$ by Wilcoxon signed rank test for the comparison of cross-reactive (green) fractions and of mimotope-reactive (red) fractions, respectively, between GRP78-298R and -502X (pooled healthy donors and donors with type 1 diabetes). T1D, type 1 diabetes.

detected by RNAseq in MHC-II^{hi} mTECs isolated from wild-type (WT) C57BL/6 mice (Fig. 6*E*) but significantly diminished or absent in $Aire^{-/-}$ mice, suggesting an *Aire*-dependent/enhanced expression. As in islets, the predominant mTEC isoform was *Padi2*, which was less impacted by *Aire* knockout in C57BL/6 mice. The enrichment of *Padi2* and *Padi4* in MHC-II^{hi} mTECs from WT versus $Aire^{-/-}$ NOD mice (Supplementary Fig. 6*A*) further supported an *Aire*-dependent/enhanced expression. No modulation by *Aire* was detected for *Grp78* or the tissue transglutaminase gene *Tgm2* responsible for the PTM transglutamination. The Assay for Transposable-Accessible Chromatin with high-throughput sequencing (ATAC-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) profiles of *Padi* genes (Supplementary Fig. 7) were also in line with an Aire-dependent expression

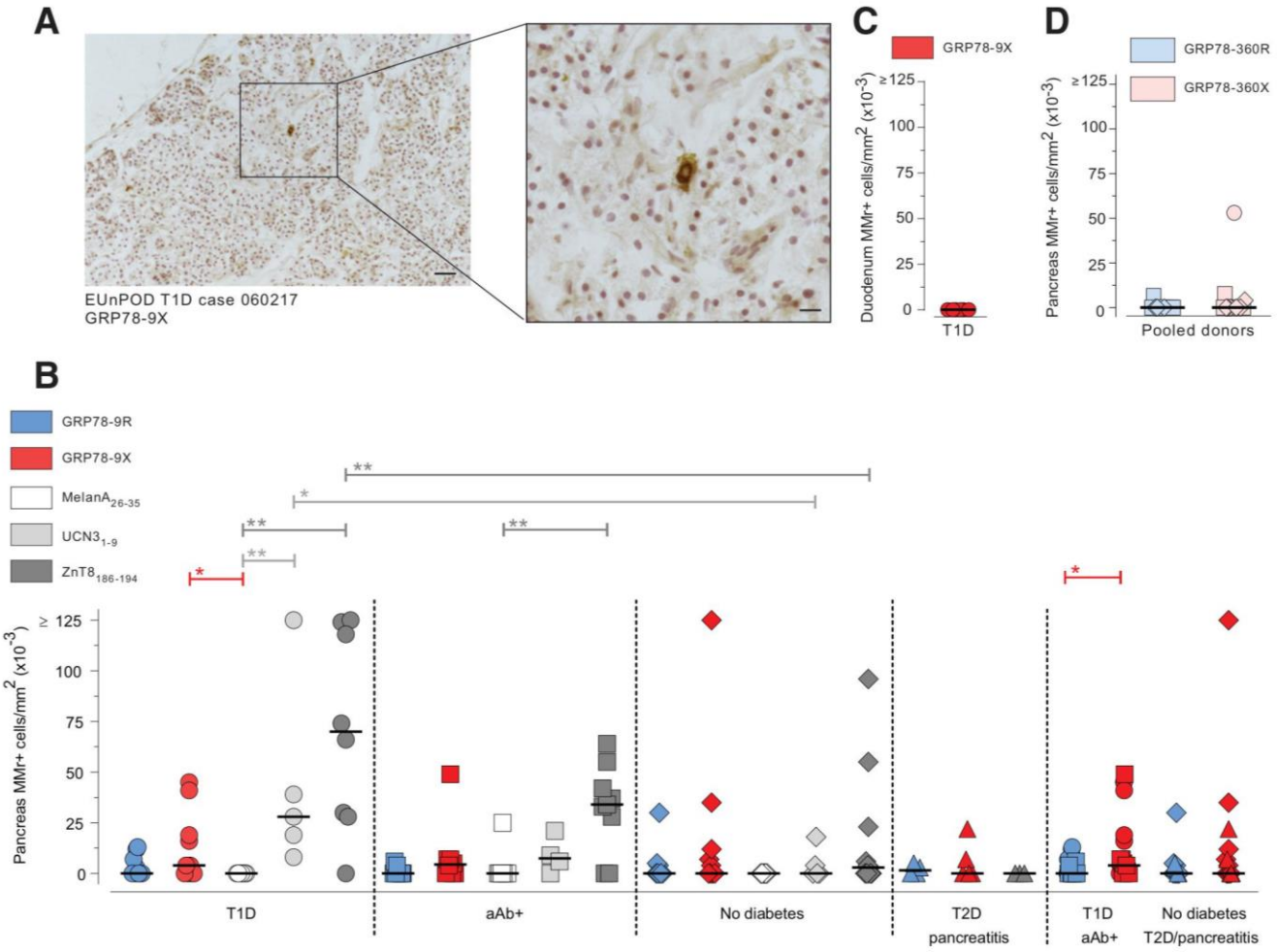


Figure 5—Pancreas-infiltrating CD8$^+$ T cells preferentially recognize the citrullinated GRP78-9X epitope. Pancreas sections from nPOD cases (Supplementary Table 2) were immunohistochemically stained in situ with MMrs loaded with the indicated GRP78 peptides, with a negative control melanocyte MelanA$_{26–35}$ peptide, and with positive control UCN3$_{1–9}$ and ZnT8$_{186–194}$ islet epitopes (low and high reactivity, respectively). *A*: Representative staining with GRP78-9X MMrs (scale bar 100 μm) and higher magnification of the dotted area (scale bar 33 μm). *B–D*: Number of MMr$^+$ cells/mm^2 section area of pancreas (*B* and *D*) and duodenal mucosa (*C*). Each point represents an individual case; bars indicate median values. T1D, type 1 diabetes; T2D, type 2 diabetes. *$P < 0.05$, **$P \leq 0.01$ by Mann-Whitney *U* test.

mechanism at an epigenetic level (24), i.e., closed transcription start sites, near absence of activating chromatin modifications (H3K4Ac, H3K9Ac), and presence of repressive modifications (H3K27me3).

Further analysis of gene sets specific for modifying enzymes (Supplementary Table 3) revealed that, among the 22 enzymatic PTMs analyzed, citrullination was the only one under the control of Aire (Fig. 6*F*; Supplementary Fig. 6*B* for NOD mice) and enriched in mature MHC-II^{high} versus immature MHC-II^{low} mTECs of WT mice (Supplementary Fig. 8). Finally, reduced Padi2 expression in the mTECs of NOD and *Aire*$^{-/-}$ C57BL/6 mice compared with WT C57BL/6 mice was confirmed at the protein level with confocal microscopy (Fig. 7).

Collectively, these observations suggest that Padi gene expression and enzymatic activity may catalyze citrullination not only in the islets but also in the thymus. In the NOD mice, the increased, age-dependent expression of *Padi2* in the islets may augment the availability of citrullinated peptides for T cell recognition. On the other hand, the reduced *Padi2* and *Padi4* expression in the mature MHC-II^{high} mTECs of NOD mice may favor the escape of T cells reactive to citrullinated peptides.

DISCUSSION

It is currently posited that posttranslationally modified epitopes may be more prone to loss of tolerance (8). This is based on the assumption that these epitopes may not be presented in the thymus, as PTMs are usually induced in a tissue-specific fashion and upregulated by inflammation. This assumption has been experimentally verified in elegant mouse models for glycosylated collagen type II (25), which is expressed in cartilages (including those from healthy donors) but not in the thymus. By analyzing GRP78 citrullination, we here document that this tenet

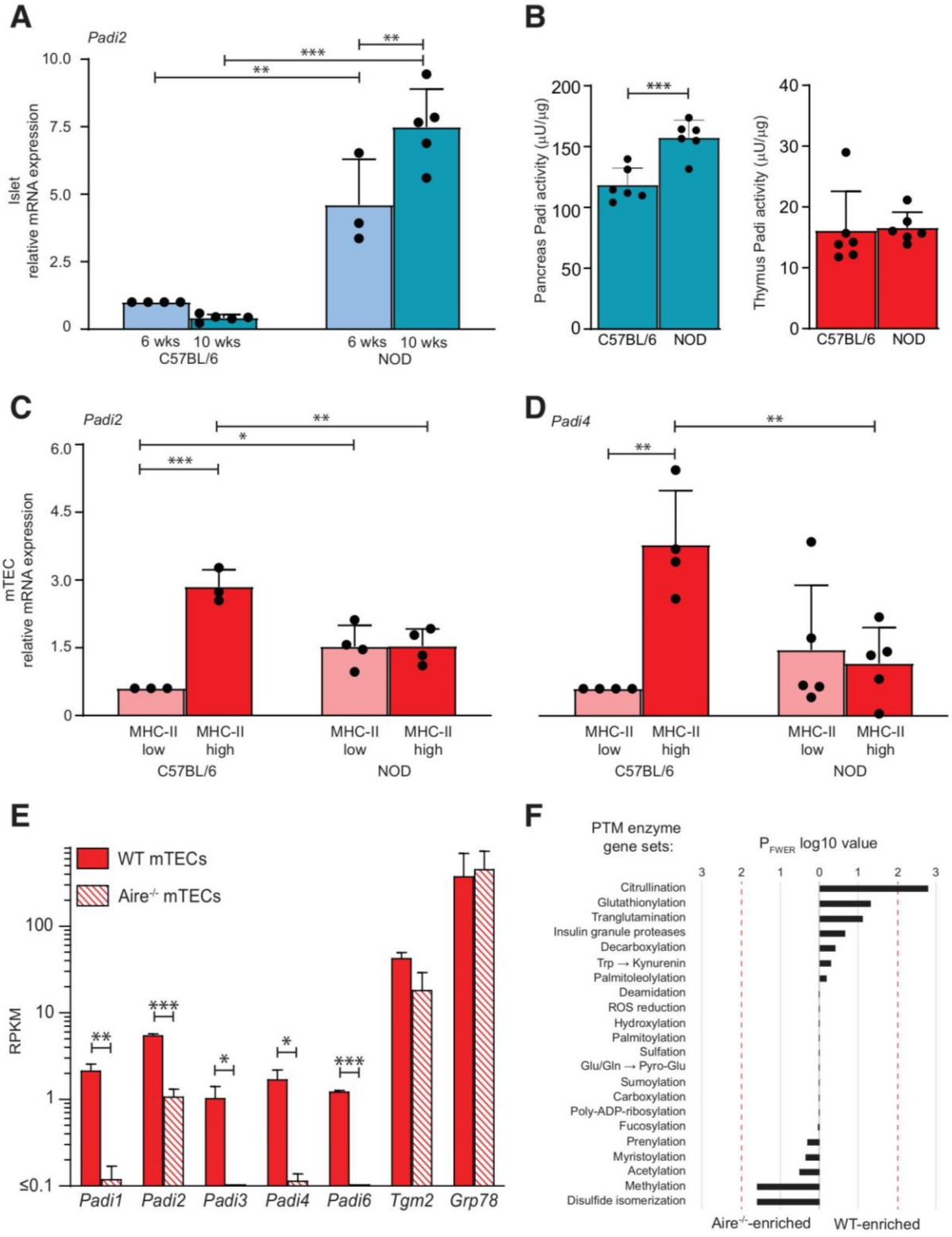


Figure 6— Gene expression of Padi citrullinating enzymes in murine mTECs and islets. *A*, *C*, and *D*: RT-qPCR expression (mean ± SD) of *Padi2* in islets from 6-week-old and 10-week-old mice (*A*) and of *Padi2* (*C*) and *Padi4* (*D*) in mTECs from 6- to 8-week-old C57BL/6 and NOD mice; $*P \leq 0.03$, $**P \leq 0.005$, $***P < 0.001$ by one-way ANOVA. *B*: Padi enzymatic activity (mean ± SD) in whole pancreata and thymi of 6-week-old C57BL/6 and NOD mice; $***P < 0.001$ by Student *t* test. *E*: Expression (median ± range) of *Padi* isoforms in WT and *Aire*$^{-/-}$ C57BL/6 mice. Murine mature MHC-II^{high} mTECs were prepared from 4- to 6-week-old mice ($n = 3$/each) and analyzed by RNAseq; $*P \leq 0.03$, $**P \leq 0.003$, $***P < 0.001$ by Welch *t* test. *F*: Enrichment of PTM enzyme gene sets in MHC-II^{high} mTECs from WT vs. *Aire*$^{-/-}$ mice. Red lines indicate the 0.01 family-wise error rate *P* (P_{FWER}) cutoff value for statistical significance (for citrullination, $P_{FWER} = 0.0016$, corresponding to a WT-normalized enrichment score of 2.4). ROS, reactive oxygen species; RPKM, reads per kilobase per million mapped reads; wks, weeks.

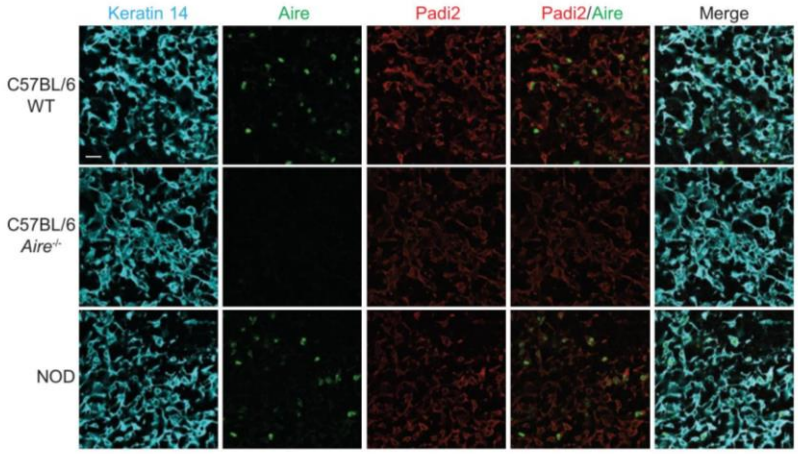


Figure 7— Padi2 protein expression in murine mTECs. Confocal microscopy of frozen thymic tissue section from 6- to 8-week-old C57BL/6 WT and *Aire*$^{-/-}$ mice and NOD mice (scale bar 20 μm). mTECs are visualized by keratin-14 staining.

may not always apply. Indeed, either native or citrullinated epitopes were preferentially recognized by circulating CD8$^+$ T cells. Moreover, this recognition was not promiscuous, since most T cells recognized exclusively one epitope version or the other.

These observations prompt two questions. First, how can GRP78 be recognized in its native form, despite its ubiquitous expression? We suggest that one driver for such recognition may be cross-reactivity with gut bacterial sequences, as mimotopes—and even homotopes—were detected for some GRP78 peptides preferred in their R version but not for those preferentially recognized as citrullinated. Of note, these R-GRP78 versions were also those displaying significant fractions of antigen-experienced CD8$^+$ T cells in some individuals and, for those with an identical bacterial match (GRP78-360R and -435R), particularly high frequencies. The final outcome of this bacterial cross recognition is likely to be tolerance, which may, however, be lost with gut dysbiosis, as recently described for IGRP microbial mimotopes (26,27).

Second, why are citrullinated GRP78 peptides not universally preferred by CD8$^+$ T cells? This may reflect the presentation of citrullinated peptides in the thymus, as both murine and human mTECs robustly expressed citrullinating Padi enzymes, most notably Padi2. *Padi* expression in murine mTECs has previously been reported along with the presence of citrullination (28). The finding of Padi2 expression and Padi enzymatic activity in the murine thymus and of citrullinated proteins in purified mTECs provides direct evidence that this gene expression can translate into PTMs. We further show that *Padi* expression is under the control of Aire in murine mTECs. The effect of Aire on promiscuous antigen expression may thus extend to citrullinated proteins. Hence, CD8$^+$ T cells recognizing native and citrullinated peptides may undergo a similar, albeit incomplete (29), thymic deletion. Interestingly, NOD mice displayed lower *Padi2* and *Padi4* expression in mature mTECs and higher expression in islets (likely in both β-cells and immune infiltrates [30]), which may exert a synergistic effect on the loss of central and peripheral tolerance against citrullinated epitopes, respectively. In line with this possibility, we previously showed that spleen T cells from NOD mice, but not from C57BL/6 mice, recognize the citrullinated version of the Grp78 protein but not its native version (9). Moreover, the lack of *Padi2*/*Padi4* upregulation between MHC-II^{lo} (and Aire$^-$) and MHC-II^{hi} (Aire$^+$) mTECs in NOD mice suggests that the Aire-driven control is impaired in this strain. The NOD thymus also harbored fewer MHC-II^{hi} mTECs, which may further contribute to defective central tolerance. Although interspecies differences are possible, it will be relevant to define whether an equivalent of this *Padi2* expression pattern in the mTECs and islets of NOD mice exists in patients with type 1 diabetes.

The finding of a similar frequency of GRP78-reactive CD8$^+$ T cells in the blood of individuals with type 1 diabetes and healthy individuals mirrors previous observations with other islet epitopes (2–4,18,20). These observations are likely to reflect the sequestration of the fraction engaged in the disease in the pancreas, where the same T cells were enriched in donors with type 1 diabetes versus nondiabetic donors (2,3). Given the notion that citrullination is enhanced in the inflammatory milieu of insulitis (30), we hypothesized that X-GRP78–reactive T cells may preferentially accumulate in the type 1 diabetic pancreas. However, despite the fact that the native GRP78-9R peptide recognized by circulating CD8$^+$ T cells was preferred in its citrullinated version in the pancreas, these T cells were not specific for type 1 diabetes.

This notion of tissue but not disease specificity echoes our recent report documenting that citrullination is also detectable in noninfiltrated islets of nonautoimmune C57BL/6 mice (30). It is possible that citrullination, although upregulated by inflammation (9,10), may be, to

a certain extent, physiological in the pancreas, as is the case in other tissues (31) and for other PTMs (25), e.g., iodinated thyroglobulin and glycosylated collagen. At the steady state, the outcome of this physiological citrullination may be peripheral tolerance, which may be broken under conditions of inflammation and ER stress that also upregulate GRP78 expression (11). Whether citrullination results in the presentation of citrullinated peptides by HLA class I molecules on β-cells or antigen-presenting cells remains an open question. Despite the presence of GRP78 peptides, immunopeptidomics studies on human and mouse islets (3,32) did not retrieve citrullinated peptides. It is noteworthy that, despite the notion of protein citrullination in inflamed cartilages (10,30,33), direct evidence of HLA-eluted citrullinated peptides is missing even for rheumatoid arthritis. While this may reflect a technical limitation of immunopeptidomics strategies, it invites the question of whether the diabetogenic role of citrullination (30) relies on T cells recognizing X-GRP78 peptides. We documented that Padi inhibition by BB-Cl-amidine treatment protects NOD mice from diabetes and is associated with decreased citrullination of GRP78 and other proteins in the pancreas, reduced anti–X-GRP78 aAb titers, and diminished formation of neutrophil extracellular traps (30). Interestingly, diabetes protection was observed upon treatment from 8 weeks of age, when insulitis is already present, suggesting an effect on disease amplification rather than priming. This is consistent with the higher titers of aAb against X-GRP78 detected in long-standing patients with type 1 diabetes (10). As citrullination is involved in the control of several biological processes (31), both physiological and pathological (e.g., neutrophil extracellular trap formation [34]), pathogenic effects independent of antigen presentation and T-cell recognition are possible. For instance, circulating aAbs against native or citrullinated GRP78 are described in patients with cancers (35), rheumatoid arthritis (36), or atherosclerosis (37). aAb binding to surface GRP78 activates endothelial cells (37) and monocytes/macrophages (36), upregulating inflammatory cytokines and adhesion molecules. Whether a similar process takes place in β-cells in the course of type 1 diabetes deserves further investigation.

This study carries some limitations. First, we did not provide evidence of natural processing and presentation of the GRP78 epitopes studied. Most of them are in close proximity to dibasic cleavage sites, which may favor their processing by the proteasome and other proteases. For the GRP78-502 epitope, which is exclusively recognized as citrullinated, we reported direct proteomic evidence for citrullination at position 510R in cytokine-treated human islets (10) and rat INS1E β-cells (9). Interestingly, this peptide overlaps with the citrullinated $GRP78_{498–512}$ epitope targeted by circulating $CD4^+$ T cells that are enriched in donors with type 1 diabetes (10). Second, the limited availability of human tissues allowed us to verify the absence of GRP78-9X–reactive $CD8^+$ T cells outside the pancreas only in the duodenum.

In conclusion, our study shows that GRP78 is targeted by circulating $CD8^+$ T cells that do not always prefer citrullinated peptides. Their recognition pattern may be shaped by cross-reactivity with gut microbial mimotopes and by thymic PADI expression. These findings mitigate the current claims that citrullinated epitopes are invariably more immunogenic than their native counterparts.

Acknowledgments. The authors thank C. Maillard, Cochin Institute, Paris, France, for technical assistance, the Cochin Institute CYBIO and KU Leuven Flow Cytometry Core Facilities for assistance with cell analysis and sorting, and J. Perez-Hernandez, Cochin Institute, Paris, France, for critical reviewing of the manuscript.

Funding. This work was supported by JDRF grants 2-SRA-2016-164-Q-R (to R.M.), 2-SRA-2015-52-Q-R (to C.M. and L.O.), and 2-SRA-2018-480-S-B (to M.N.) and postdoctoral fellowship 3-PDF-2020-942-A-N (to Z.Z.); The Leona M. and Harry B. Helmsley Charitable Trust (1901-03689); Fondation Francophone pour la Recherche sur le Diabète; the European Foundation for the Study of Diabetes (EFSD/JDRF/Lilly European Programme in Type 1 Diabetes Research 2015); Agence Nationale de la Recherche (ANR-19-CE15-0014-01) and Fondation pour la Recherche Medicale (EQU20193007831) (to R.M.); and the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK099317, R01DK032083 to M.N.). M.B. and A.Cal. were supported by a fellowship from the Flemish Research Foundation (12R0719N and 1189518N, respectively). F.D., D.L.E., C.M., L.O., and R.M. received funding from the Innovative Medicines Initiative 2 Joint Undertaking under grant agreements 115797 and 945268 (INNODIA and INNODIA HARVEST), which receive support from the EU Horizon 2020 program, JDRF, and The Leona M. & Harry B. Helmsley Charitable Trust. This research was performed with the support of the nPOD (RRID SCR_014641), a collaborative type 1 diabetes research project supported by JDRF (nPOD: 5-SRA-2018-557-Q-R) and The Leona M. & Harry B. Helmsley Charitable Trust (grant 2018PG-T1D053). Organ procurement organizations (OPO) partnering with nPOD to provide research resources are listed at www.jdrfnpod.org/for-partners/npod-partners. The content and views expressed are the responsibility of the authors and do not necessarily reflect an official view of nPOD.

Duality of Interest. INNODIA and INNODIA HARVEST also receive support from the European Federation of Pharmaceutical Industries and Associations. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. F.S., M.G., C.M., L.O., and R.M. contributed to study conceptualization. M.E.A., F.S., M.B., L.N., N.B., A.Cal., M.G., M.I., A.I.L., A.Car., Z.Z., B.B., M.L.C., G.S., F.D., M.N., D.L.E., S.Y., S.P., M.J.M., Y.V., J.V., S.B., C.M., L.O., and R.M. contributed to development of study methodology. M.E.A., F.S., M.B., L.N., N.B., A.Cal., M.G., M.I., A.I.L., A.Car., G.A., Z.Z., B.B., M.L.C., G.S., M.N., L.O., and R.M. contributed to data generation. M.E.A., F.S., M.B., L.N., N.B., A.Cal., M.G., A.Car., M.L.C., G.S., M.J.M., L.O., and R.M. contributed to data curation. M.E.A., F.S., M.B., M.G., L.O., and R.M. contributed to writing the manuscript. M.E.A., F.S., M.B., L.N., N.B., A.Cal., M.G., M.I., L.O., and R.M. contributed to data visualization. L.O. and R.M. contributed to study supervision. S.Y., C.M., L.O., and R.M. contributed to funding acquisition. R.M. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the nPOD 12th Annual Scientific Meeting, Tampa, FL, 23–26 February 2020.

References

1. Carré A, Richardson SJ, Larger E, Mallone R. Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? Diabetologia 2021;64:15–25
2. Culina S, Lalanne AI, Afonso G, et al.; ImMaDiab Study Group. Islet-reactive $CD8^+$ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013
3. Gonzalez-Duque S, Azoury ME, Colli ML, et al. Conventional and neo-antigenic peptides presented by β cells are targeted by circulating naïve CD8+ T cells in type 1 diabetic and healthy donors. Cell Metab 2018;28:946–960.e6
4. Azoury ME, Tarayrah M, Afonso G, et al. Peptides derived from insulin granule proteins are targeted by $CD8^+$ T cells across MHC class I restrictions in humans and NOD mice. Diabetes 2020;69:2678–2690
5. Alanio C, Lemaitre F, Law HK, Hasan M, Albert ML. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. Blood 2010;115:3718–3725
6. Yu W, Jiang N, Ebert PJ, et al. Clonal deletion prunes but does not eliminate self-specific αβ CD8(+) T lymphocytes. Immunity 2015;42:929–941
7. Mallone R, Eizirik DL. Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? Diabetologia 2020;63:1999–2006
8. James EA, Pietropaolo M, Mamula MJ. Immune recognition of β-cells: neoepitopes as key players in the loss of tolerance. Diabetes 2018;67:1035–1042
9. Rondas D, Crèvecoeur I, D'Hertog W, et al. Citrullinated glucose-regulated protein 78 is an autoantigen in type 1 diabetes. Diabetes 2015;64:573–586
10. Buitinga M, Callebaut A, Marques Câmara Sodré F, et al. Inflammation-induced citrullinated glucose-regulated protein 78 elicits immune responses in human type 1 diabetes. Diabetes 2018;67:2337–2348
11. Wang J, Lee J, Liem D, Ping P. HSPA5 gene encoding Hsp70 chaperone BiP in the endoplasmic reticulum. Gene 2017;618:14–23
12. Vig S, Buitinga M, Rondas D, et al. Cytokine-induced translocation of GRP78 to the plasma membrane triggers a pro-apoptotic feedback loop in pancreatic beta cells. Cell Death Dis 2019;10:309
13. Babon JA, DeNicola ME, Blodgett DM, et al. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. Nat Med 2016;22:1482–1487
14. Leisner C, Loeth N, Lamberth K, et al. One-pot, mix-and-read peptide-MHC tetramers. PLoS One 2008;3:e1678
15. Casteels K, Waer M, Laureys J, et al. Prevention of autoimmune destruction of syngeneic islet grafts in spontaneously diabetic nonobese diabetic mice by a combination of a vitamin D3 analog and cyclosporine. Transplantation 1998;65:1225–1232
16. Callebaut A, Derua R, Vig S, Delong T, Mathieu C, Overbergh L. Identification of deamidated peptides in cytokine-exposed MIN6 cells through LC-MS/MS using a shortened digestion time and inspection of MS2 spectra. J Proteome Res 2021;20:1405–1414
17. Guyon C, Jmari N, Padonou F, et al. Aire-dependent genes undergo Clp1-mediated 3'UTR shortening associated with higher transcript stability in the thymus. eLife 2020;9:e52985
18. Wiedeman AE, Muir VS, Rosasco MG, et al. Autoreactive CD8+ T cell exhaustion distinguishes subjects with slow type 1 diabetes progression. J Clin Invest 2020;130:480–490
19. Yeo L, Pujol-Autonell I, Baptista R, et al. Circulating β cell-specific $CD8^+$ T cells restricted by high-risk HLA class I molecules show antigen experience in children with and at risk of type 1 diabetes. Clin Exp Immunol 2020;199:263–277
20. Skowera A, Ladell K, McLaren JE, et al. β-Cell-specific CD8 T cell phenotype in type 1 diabetes reflects chronic autoantigen exposure. Diabetes 2015;64:916–925
21. Genest O, Wickner S, Doyle SM. Hsp90 and Hsp70 chaperones: collaborators in protein remodeling. J Biol Chem 2019;294:2109–2120
22. Bender C, Rodriguez-Calvo T, Amirian N, Coppieters KT, von Herrath MG. The healthy exocrine pancreas contains preproinsulin-specific CD8 T cells that attack islets in type 1 diabetes. Sci Adv 2020;6:eabc5586
23. Crèvecoeur I, Gudmundsdottir V, Vig S, et al. Early differences in islets from prediabetic NOD mice: combined microarray and proteomic analysis. Diabetologia 2017;60:475–489
24. Handel AE, Shikama-Dorn N, Zhanybekova S, et al. Comprehensively profiling the chromatin architecture of Tissue restricted antigen expression in thymic epithelial cells over development. Front Immunol 2018;9:2120
25. Raposo B, Merky P, Lundqvist C, et al. T cells specific for post-translational modifications escape intrathymic tolerance induction. Nat Commun 2018;9:353
26. Hebbandi Nanjundappa R, Ronchi F, Wang J, et al. A gut microbial mimic that hijacks diabetogenic autoreactivity to suppress colitis. Cell 2017;171:655–667.e17
27. Tai N, Peng J, Liu F, et al. Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. J Exp Med 2016;213:2129–2146
28. Engelmann R, Biemelt A, Cordshagen A, Johl A, Kuthning D, Müller-Hilke B. The prerequisites for central tolerance induction against citrullinated proteins in the mouse. PLoS One 2016;11:e0158773
29. Davis MM. Not-So-negative selection. Immunity 2015;43:833–835
30. Sodré FMC, Bissenova S, Bruggeman Y, et al. Peptidylarginine deiminase inhibition prevents diabetes development in NOD mice. Diabetes 2021;70:516–528
31. Alghamdi M, Alasmari D, Assiri A, et al. An overview of the intrinsic role of citrullination in autoimmune disorders. J Immunol Res 2019;2019:7592851
32. Wan X, Vomund AN, Peterson OJ, Chervonsky AV, Lichti CF, Unanue ER. The MHC-II peptidome of pancreatic islets identifies key features of autoimmune peptides. Nat Immunol 2020;21:455–463
33. Tilvawala R, Nguyen SH, Maurais AJ, et al. The rheumatoid arthritis-associated citrullinome. Cell Chem Biol 2018;25:691–704.e6
34. Papayannopoulos V. Neutrophil extracellular traps in immunity and disease. Nat Rev Immunol 2018;18:134–147
35. Mintz PJ, Kim J, Do KA, et al. Fingerprinting the circulating repertoire of antibodies from cancer patients. Nat Biotechnol 2003;21:57–63
36. Lu MC, Lai NS, Yu HC, Huang HB, Hsieh SC, Yu CL. Anti-citrullinated protein antibodies bind surface-expressed citrullinated Grp78 on monocyte/macrophages and stimulate tumor necrosis factor alpha production. Arthritis Rheum 2010;62:1213–1223
37. Crane ED, Al-Hashimi AA, Chen J, et al. Anti-GRP78 autoantibodies induce endothelial cell activation and accelerate the development of atherosclerotic lesions. JCI Insight 2018;3:e99363

SUPPLEMENTARY MATERIAL

CD8$^+$ T cells variably recognize native versus citrullinated GRP78 epitopes in type 1 diabetes

Marie Eliane Azoury[1*], Fatoumata Samassa[1*], Mijke Buitinga[2], Laura Nigi[3], Noemi Brusco[3], Aïsha Callebaut[2], Matthieu Giraud[4], Magali Irla[5], Ana Ines Lalanne[1], Alexia Carré[1], Georgia Afonso[1], Zhicheng Zhou[1], Barbara Brandao[1], Maikel L. Colli[6], Guido Sebastiani[3], Francesco Dotta[3], Maki Nakayama[7], Decio L. Eizirik[6,8], Sylvaine You[1], Sheena Pinto[9], Mark J. Mamula[10], Yann Verdier[11], Joelle Vinh[11], Soren Buus[12], Chantal Mathieu[2], Lut Overbergh[2] Roberto Mallone [1,13]

[1]Université de Paris, Institut Cochin, CNRS, INSERM, 75014 Paris, France.
[2]KU Leuven, Laboratory of Clinical and Experimental Endocrinology, 3000 Leuven, Belgium.
[3]University of Siena, Department of Medicine, Surgery and Neuroscience, Diabetes Unit and Fondazione Umberto di Mario ONLUS, Toscana Life Sciences, 53100 Siena, Italy.
[4]Université de Nantes, INSERM UMR1064, Centre de Recherche en Transplantation et Immunologie (CRTI), 44093 Nantes, France.
[5]Aix-Marseille University, CNRS, INSERM, CIML, Centre d'Immunologie de Marseille-Luminy, Marseille, France[6]Université Libre de Bruxelles, Center for Diabetes Research and Welbio, Medical Faculty, 1070 Brussels, Belgium.
[7]Barbara Davis Center for Childhood Diabetes, University of Colorado School of Medicine, Aurora, CO 80045, USA.
[8]Indiana Biosciences Research Institute, Indianapolis, IN 46202, USA.
[9]Deutsches Krebsforschungszentrum (DKFZ), Division of Developmental Immunology, 69120 Heidelberg, Germany.
[10]Yale University School of Medicine, New Haven, CT 06510, USA.
[11]ESPCI Paris, PSL University, Spectrométrie de Masse Biologique et Protéomique, CNRS USR3149, 75005 Paris, France.
[12]Panum Institute, Department of International Health, Immunology and Microbiology, D-2200 Copenhagen, Denmark.
[13]Assistance Publique Hôpitaux de Paris, Service de Diabétologie et Immunologie Clinique, Cochin Hospital, 75014 Paris, France.

*M.E.A. and F.S. contributed equally to this work.

SUPPLEMENTARY METHODS

mTEC isolation

For RT-qPCR and proteomics, thymic lobes were harvested from 6-8-week-old C57BL/6 mice and NOD mice, cleaned of connective tissue and fat, cut into fragments, washed in RPMI and gently agitated thrice for 30 s to release excess thymocytes. At each wash, the fragments were allowed to settle for 10 min on ice, followed by removal of the thymocyte-rich supernatant. Fragments were subsequently digested for 20 min at 37°C in ultra-low attachment 6-well plates (Corning) using 5 ml of ice-cold Liberase TH (2.5 U/mL; Roche) supplemented with 100 U/ml DNase I (Applichem). Supernatants were then removed and collected in ice-cold PBS supplemented with 0.5% BSA and 2mM EDTA. This enzymatic digestion was repeated twice for 15 min with 2.5 ml Liberase. Syringe needles (18G and 25G) were used to break up tissue after the second and third digestion, respectively. Cells were filtered through a 100-µm mesh before enrichment for TECs. To this end, $2x10^8$ cells/ml were incubated in PBS/2% FCS with 2.6 µg/ml anti-CD90.2 mAb (clone 30-H12, Sigma) on a rotating shaker for 30 min at 4°C, washed and plated ($1x10^7$ cells/ml) on panning plates previously coated overnight with 10 µg/ml goat anti-rat Ab (Sigma) in 50 mM Tris buffer pH 9.5. After 30 min at room temperature, cells were collected, washed and incubated with the Fixable Viability Dye eFluor780 (eBioscience). Cells ($40x10^6$/ml) were then stained for CD45-PE/Cy7 (RRID:AB_2734986), MHC Class II (MHC-II)-AF488 (RRID:AB_493138 and RRID:AB_493147 for C57BL/6 and NOD mice, respectively), EPCAM-BV421 (RRID:AB_2563983), BP-1-APC (RRID:AB_2762698), *Ulex europaeus* agglutinin-1 (UEA-1-biotin; Vector Laboratories), and streptavidin-PE (eBioscience). Cells ($40x10^6$/ml) were sorted with a BD Influx (100-µm nozzle, 18 psi) equipped with 488 nm (200 mW), 640 nm (120 mW), 405 nm (100 mW) and 561 nm (150 mW) lasers. Sorted cells were collected in 1.5-ml tubes in PBS/2% FCS.

For RNAseq, mouse thymi were extracted from a pool of 4 C57BL/6 mice, cut into small pieces, agitated into RPMI medium to release thymocytes, and digested with collagenase D (1 mg/ml; Roche) and DNase I (1 mg/ml; Sigma) for 30 min at 37 °C. The remaining fragments were digested with collagenase/dispase (2mg/ml; Roche) and DNase I (2 mg/ml) at 37 °C to obtain a cell suspension. After filtration through a 70-µm mesh, cells were resuspended in PBS containing 1% fetal bovine serum (FBS) and 5 mM EDTA. The remaining thymocytes were removed by magnetic depletion with CD45 MicroBeads (Miltenyi). Cells were then stained with CD45-PerCPCy5.5 (1:50; RRID:AB_893340), Ly51-PE (1:800; RRID:AB_313365), and MHC-II I-A/I-E-APC (1:1,200; RRID:AB_469455). MHC-II$^{high/low}$ mTECs (CD45$^-$Ly51$^-$ MHC-II$^{high/low}$) were sorted into TRIzol (ThermoFisher) on a BD FACSAria III instrument.

RT-qPCR and measurement of Padi enzymatic activity

Total RNA was extracted from sorted mTECs or 50 islets, using the Single Cell RNA Purification Kit (Norgen) and cDNA synthesized with SuperScript VILO (Invitrogen). The RT-qPCR forward (F) and reverse (R) primers were: *Padi2*-F: 5'-CCGCCGGGTATGAAATAGTCC-3', *Padi2*-R: 5'-CGCCGGTGTACTTGACCAC-3'; housekeeping *Act*-F: 5'-AGAGGGAAATCGTGCGTGAC-3', *Act*-R: 5'-CAATAGTGATGACCTGGCCGT-3'; housekeeping *Rpl27*-F: 5'-GTCGAGATGGGCAAGTTCAT-3', *Rpl27*-R: 5'-TTCTTCACGATGACGGCTTT-3'; and predesigned *Hprt* TaqMan primers (ThermoFisher). RT-qPCR was performed using 4 pmol primers, 0.2 µl cDNA and 5 µl Fast SYBR Green Master Mix (Applied Biosystems) on a StepOnePlus RT-PCR System (Applied Biosystems). The relative fold gene expression was calculated using the delta-delta Ct method.

Padi activity was measured with the Ab-based assay for Padi activity (ABAP; ModiQuest Research).

Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

mTECs were washed thrice with PBS at 400g for 5 min and lysed in a buffer composed of 7 M urea, 2 M thiourea, 4% CHAPS, 40 mM Tris base, 1% dithiothreitol and Complete protease inhibitors (Roche). Debris were removed by centrifugation for 10 min at 13,000 rpm at 4°C. Samples were dialyzed and reduction, alkylation, protein precipitation, digestion and desalting performed as described (1). The purified peptides were vacuum-dried and dissolved in mobile phase A, containing 2% acetonitrile and 0.1% formic acid. An estimated final amount of 0.5 µg/µl was injected. The sample was separated by reversed-phase chromatography using a micropillar array column (µPAC^{TM} C18, 200 cm, PharmaFluidics). A linear gradient of mobile phase B (0.1% formic acid in 98% acetonitrile) from 1% to 40% in 80 min was followed by a steep increase to 100% mobile phase B in 5 min. After 5 min at 100% mobile phase B, a steep decrease to 1% mobile phase B was achieved in 5 min and 1% mobile phase B was maintained for 35 min with a flow rate of 750 nl per minute. LC-MS/MS was performed on a Q-Exactive Plus equipped with a nanospray ion source (ThermoFisher). Full-scan spectrum (350 to 1850 m/z, resolution 70,000, automatic gain control 3E6, maximum injection time 100 ms) was followed by high-energy collision-induced dissociation (HCD) tandem mass spectra with a run time of 120 min. Peptide ions were selected for fragmentation by tandem MS as the 20 most intense peaks of a full-scan mass spectrum. HCD scans were acquired in the Orbitrap

(resolution 17,500, automatic gain control 1E5, maximum injection time 80 ms). Peptides were identified by MASCOT (Matrix Science) using SwissProt (Homo sapiens, 169,779 entries) as a database via Proteome Discoverer 2.2, incorporating Percolator for peptide validation. Oxidation (M), deamidation (N/Q), and deamidation (R) (referring to citrullination), were included as variable modifications, carbamidomethylation (C) as a fixed modification. Two missed cleavages were allowed, peptide tolerance was set at 5 ppm and MS/MS tolerance at 20 mmu. MS/MS spectra were checked manually for the presence of citrullinated residues as described (2).

RNAseq and gene set enrichment analysis (GSEA)

Total RNA was extracted following the TRIzol protocol, adding GlycoBlue (ThermoFisher) as an RNA carrier. This RNA was used to generate poly-A-selected transcriptome libraries using the non-directional TruSeq V3 RNA Sample Prep Kit (Illumina) following the manufacturer's protocol. Sequencing was carried out on an Illumina HiSeq 2000 machine and was paired-end (2x100 bp) for MHC-II^{high} mTECs and single-end (50 bp) for MHC-II^{low} mTECs. These datasets have been deposited under GEO: (GSE submitted). RNAseq data from published (3) MHC-II^{high} and MHC-II^{low} mTEC (GSE140815) and from *Aire*$^{-/-}$ MHC-II^{high} mTEC (GSE140683) biological replicates were obtained with the same procedure.

Low-quality sequencing reads were removed using the Illumina CASAVA 1.8 pipeline and homogenized to 50-bp reads by trimming and merging the paired-end 100-bp reads. These reads were then aligned to the mm9 mouse reference genome using Bowtie (4). Read numbers were generated using the intersectBed with parameters -f 1.0 and the coverageBed programs of the BEDtool distribution (5), with a GTF annotation file that was generated from the UCSC Table Browser, in choosing mm9 mouse genome/Genes and Gene Predictions/RefSeq Genes/refGene. Reads per kilobase per million mapped reads (RPKM) values were then computed for each sample according to the transcript length of each gene and the global number of reads that map to the mm9 genome.

The enrichment of sets of genes driving PTMs between wild-type (WT) and *Aire*$^{-/-}$ MHC-II^{hi} mTECs and between MHC-II^{high} and MHC-II^{low} mTECs was analyzed with the GSEA software (6). To this end, the sequencing read numbers corresponding to each group of biological replicates in each comparison were normalized using DESeq (7) and used to run GSEA with a "classic" scoring scheme and n=10,000 permutations. The conservative family-wise error rate statistics (P_{FWER}) was used to determine whether a set of genes was significantly over-represented in a group of mTECs versus another.

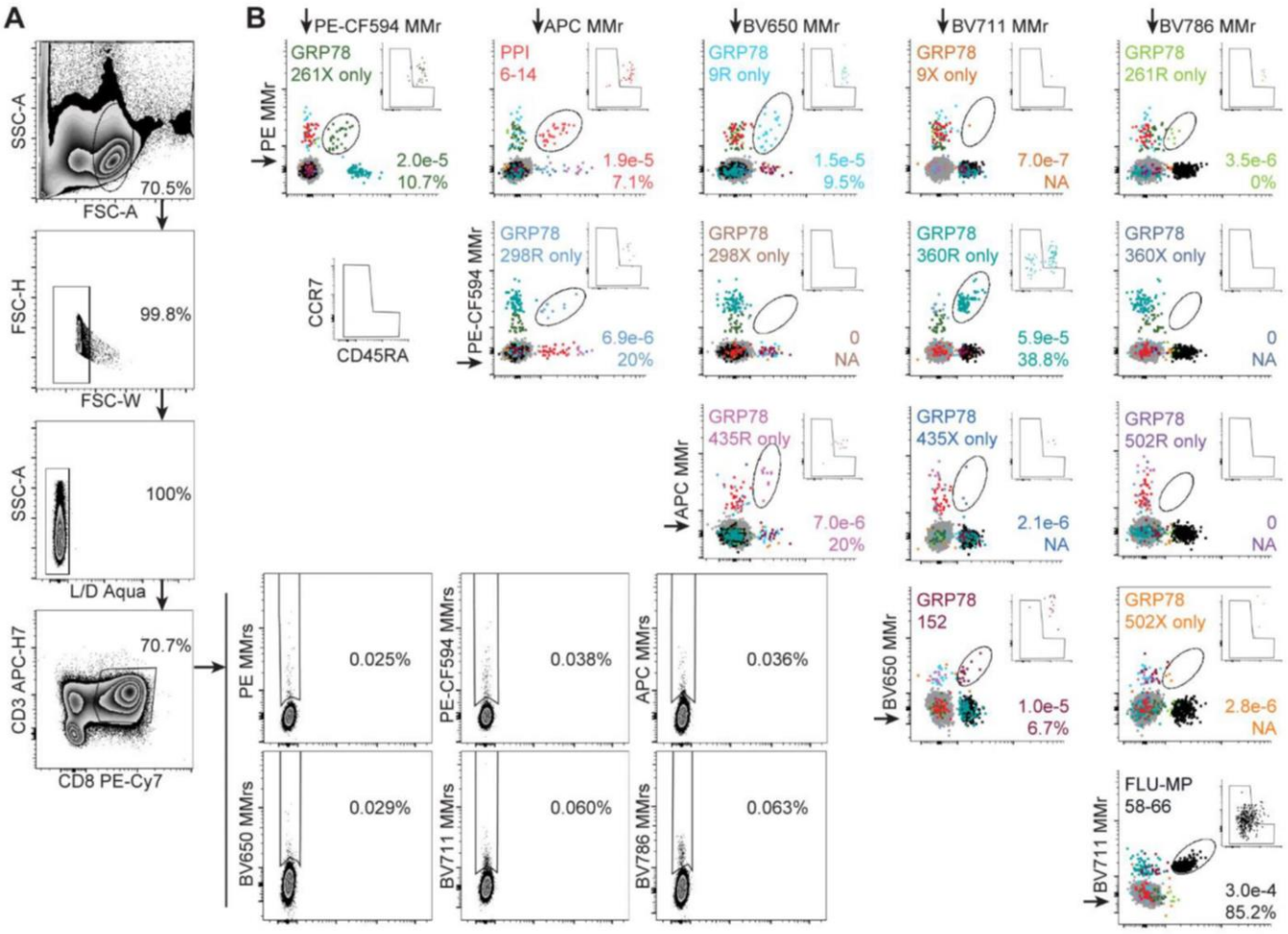
Microarray data from MHC-II^{high} mTECs isolated from 3 WT and 3 *Aire*$^{-/-}$ NOD mice were obtained from the NCBI GEO database (GSE12073) (8). CEL raw data files were processed with GenePattern (9) and the ExpressionFileCreator module for normalization using the RMA method (10) with background correction. Normalized data were plotted as fold change expression. False discovery rate *p* values were computed using the NCBI GEO2R online tool. Enrichment of PTM enzyme gene sets between WT and *Aire*$^{-/-}$ NOD MHC-II^{high} mTECs was analyzed using GSEA on normalized microarray expression data using a "classic" scoring scheme and nominal *p* values.

Confocal microscopy

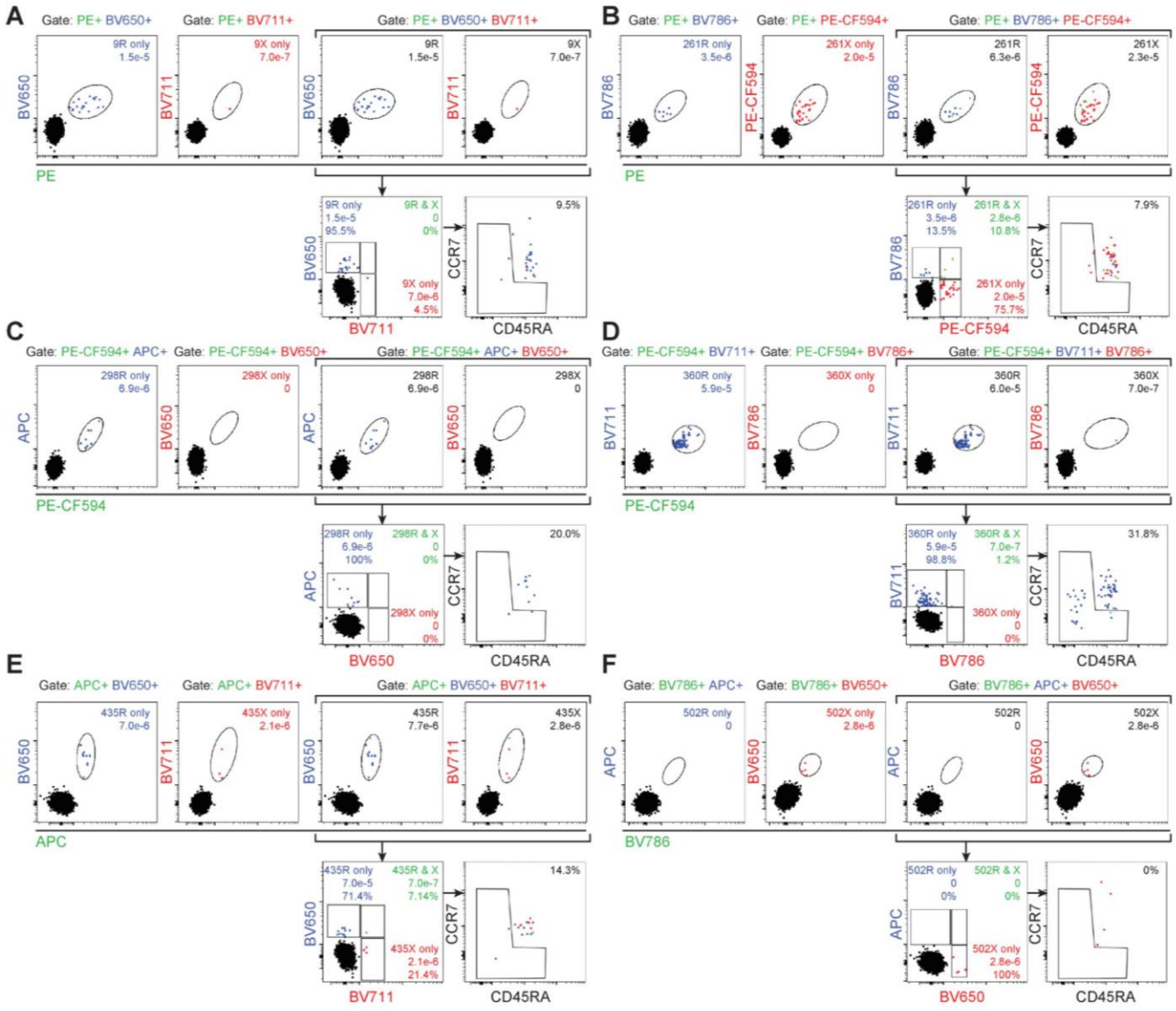
Frozen thymic tissue sections from 6-8-week-old WT and *Aire*$^{-/-}$ C57BL/6 mice and from NOD mice were fixed with 2% paraformaldehyde (Sigma) and incubated for 10 min in a saturation buffer composed of 0.1 M Tris pH 7.4, 5% mouse serum, 2% bovine serum albumin (Axday), 0.01% Triton X-100. Sections were then stained as previously described (11), with primary antibodies: rabbit anti-mouse keratin-14 (RRID:AB_2565048), mouse anti-PADI2 (RRID: AB_2881762), AF488-labeled rat anti-mouse Aire (RRID:AB_10854132); and secondary antibodies: Cy5-labeled goat anti-rabbit IgG (RRID: AB_2534032) and Cy3-labeled goat anti-mouse IgG (RRID:AB_893530). Sections were mounted with Mowiol (Calbiochem) and acquired on a LSM 780 confocal microscope (Carl Zeiss Microscopy).

Statistics

Significance was assessed with a cutoff value of α=0.05 using two-tailed Student t test or Mann-Whitney U test according to normal or non-normal distribution, respectively. Variables displaying unequal variances were analyzed with the Welch t test. Paired datasets were analyzed with the Wilcoxon signed rank test.



Supplementary Figure 1. Gating strategy for the combinatorial analysis of GRP78 MMr$^+$CD8$^+$ T cells in T1D and healthy donors. (A) Frozen-thawed PBMCs from healthy donor C59 were magnetically depleted of CD8$^-$ cells before staining, acquisition and analysis. Cells were sequentially gated on small lymphocytes, singlets, live cells (Live/Dead Aqua$^-$), CD3$^+$CD8$^+$ T cells, and total PE$^+$, PE-CF594$^+$, APC$^+$, BV650$^+$, BV711$^+$, and BV786$^+$ MMr$^+$ T cells. Using Boolean operators, these latter gates allowed selective visualization of each double-MMr$^+$ population by including only those events positive for the corresponding fluorochrome pair (see Research Design and Methods). **(B)** The final readout obtained for the 15 peptides analyzed: 12 native/citrullinated GRP78 peptides, a GRP78$_{152\text{-}161}$ (GRP78-152) peptide without R residues, and PPI$_{6\text{-}14}$ and Flu-MP$_{58\text{-}66}$ controls). Events corresponding to each epitope-reactive population are overlaid in different colors within each plot, with MMr$^-$ events overlaid in light gray. The small dot plots on the right of each panel depict CD45RA (x axis) and CCR7 (y axis) expression in the corresponding MMr$^+$ fraction. Numbers in each panel indicate the MMr$^+$CD8$^+$ T-cell frequency out of total CD8$^+$ T cells and the percent naive (CD45RA$^+$CCR7$^+$) fraction among MMr$^+$ cells. This first gating strategy only allows to visualize CD8$^+$ T cells reactive to either the R or X isoform of each GRP78 peptides, and these T cells are therefore designated as 'R only' or 'X only', respectively. This strategy was therefore further modified to differentially visualize CD8$^+$ T cells reactive to R only, X only or both peptides (see Supplemental Figure 2).

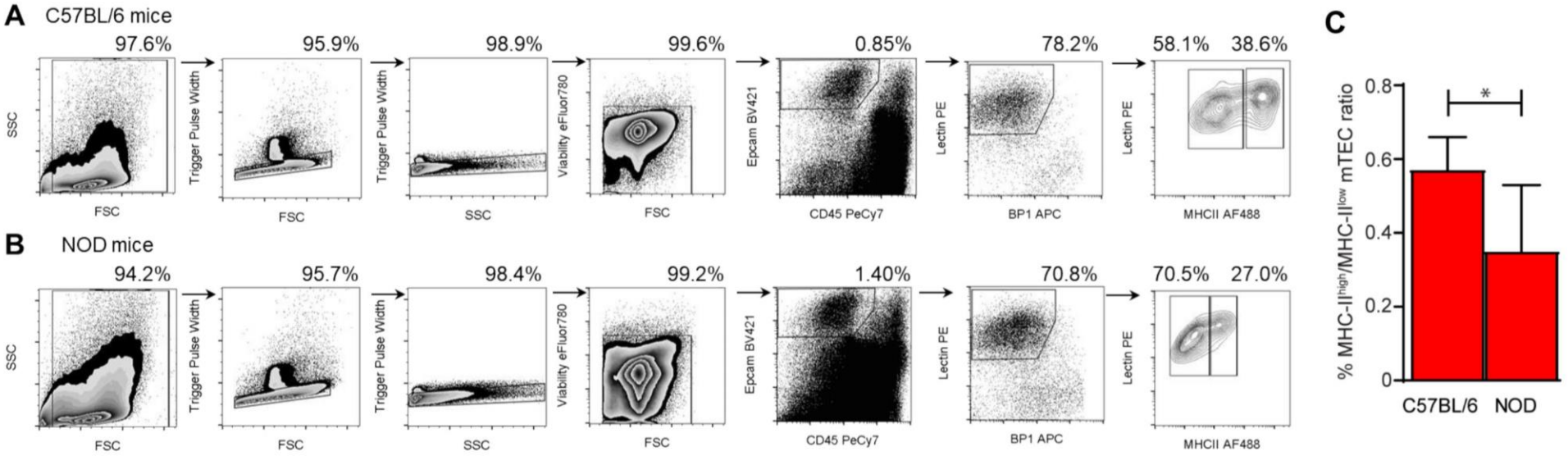


Supplementary Figure 2. Modified combinatorial gating strategy to analyze the cross-reactivity between R-GRP78 and X-GRP78 peptide isoforms. A representative example is shown for healthy donor C59. Each panel shows one of the peptides studied: GRP78-9 **(A)**, -261 **(B)**, -298 **(C)**, -360 **(D)**, -435 **(E)**, and -502 **(F)**.

First, $CD8^+$ T cells recognizing only the R-GRP78 (blue, 'R only') or X-GRP78 isoform (red, 'X only') were gated on double MMr^+ events, excluding the other four fluorochromes (first two dot plots of each panel, which are the same of Supplemental Fig. 1; MMr^- events overlaid in black). For example, in panel B '261R only' MMr^+ cells (PE^+BV786^+) were visualized by gating on $PE^+BV786^+PE\text{-}CF594^-APC^-BV650^-BV711^-$ events; and '261X only' MMr^+ cells ($PE^+PE\text{-}CF594^+$) were visualized by gating on $PE^+PE\text{-}CF594^+APC^-BV650^-BV711^-BV786^-$ events.

Second, $CD8^+$ T cells recognizing R-GRP78 or X-GRP78 peptides (i.e. including cross-reactive T cells) were gated on triple MMr^+ events, i.e. by excluding three fluorochromes but not the non-shared fluorochrome coding for the MMr loaded with the other peptide isoform (third and fourth dot plots of each panel). For example, in panel B '261R' and '261X' MMr^+ cells (i.e. cross-reactive or not; PE^+BV786^+ and PE^+PE-

$CF594^+$, respectively) were visualized by gating on $PE^+BV786^+APC^-BV650^-BV711^-$ (i.e. not excluding
CF594 as before) and $PE^+PE\text{-}CF594^+APC^-BV650^-BV711^-$ (i.e. not excluding BV786), respectively.
Third, these two dot plots are combined and visualized by plotting them for the two fluorochromes not sha
by the R and X isoform (first dot plot in the bottom line of each panel). In this final dot plot, single-posit
events correspond to the 'R only' and 'X only' fractions (blue and red, respectively), while double-posit
events correspond to the 'R and X' fraction (i.e. cross-reactive; green). Events negative for all M
fluorochromes ($PE^-PE\text{-}CF594^-APC^-BV650^-BV711^-BV786^-$) are overlaid in black in each dot plot to set
double- or triple-MMr^+ gates. The bottom right dot plots of each panel depict CD45RA and CCR7 express
in the corresponding MMr^+ fractions using the same blue, red and green color code for 'R only', 'X or
and 'R and X' MMr^+ events, respectively.

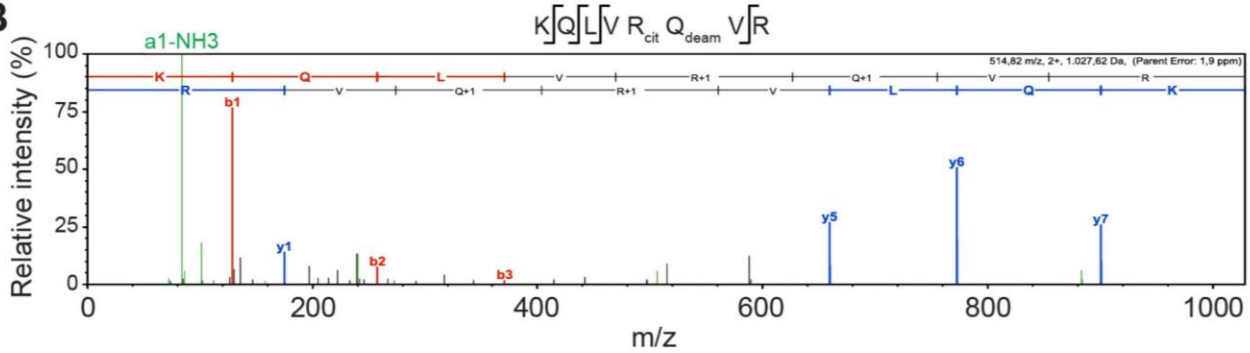


Supplementary Figure 3. Gating strategy and distribution of MHC-II^{low} and MHC-II^{high} mTECs in C57BL/6 and NOD mice. (A) mTECs from C57BL/6 mice. **(B)** mTECs from NOD mice. mTECs were gated as CD45^{-} EPCAM^{+} BP1$^{neg/low}$ Lectin (UEA-1)^{high} and MHC-II^{low} or MHC-II^{high}. **(C)** Ratio (median ± range) of MHC-II^{high}/MHC-II^{low} mTECs in 6-8-week-old C57BL/6 (n=4) and NOD mice (n=5); *p=0.032 by Mann-Whitney U test.

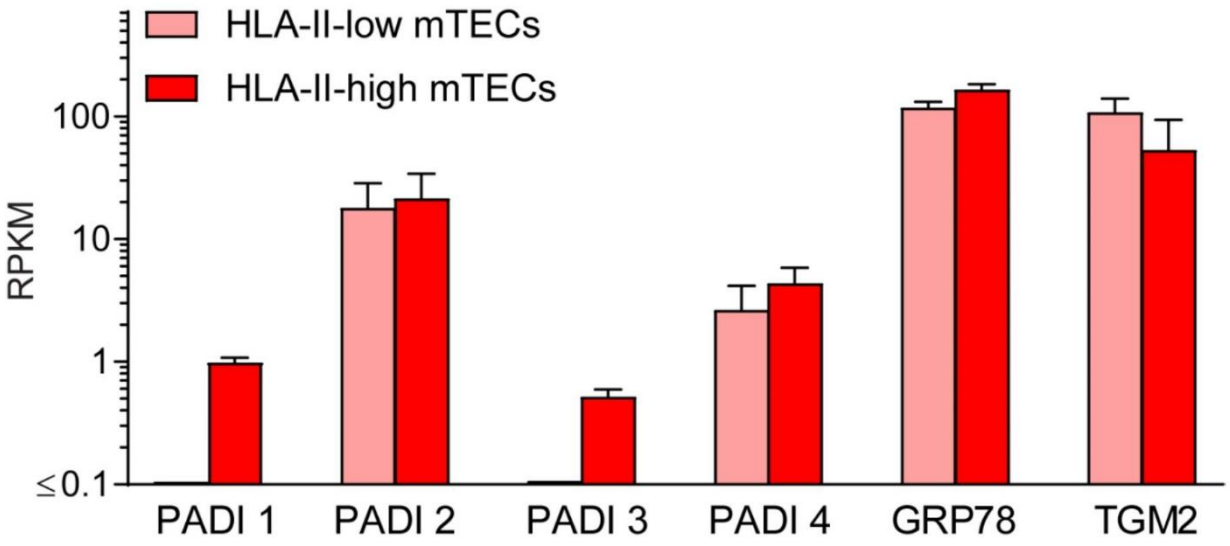
A

Name	Sequence	Modifications
EF-hand domain-containing protein 1	NGYAVVRRPTMGIGGDR	1xCitrullinated [R7]; 1xOxidation [M11]
Hepatocyte growth factor-regulated tyrosine kinase substrate	AAEEAERQR	1xCitrullinated [R7]; 1xDeamidated [Q8]
Myosin-9	KQLVRQVR	1xCitrullinated [R5]1xDeamidated [Q6]

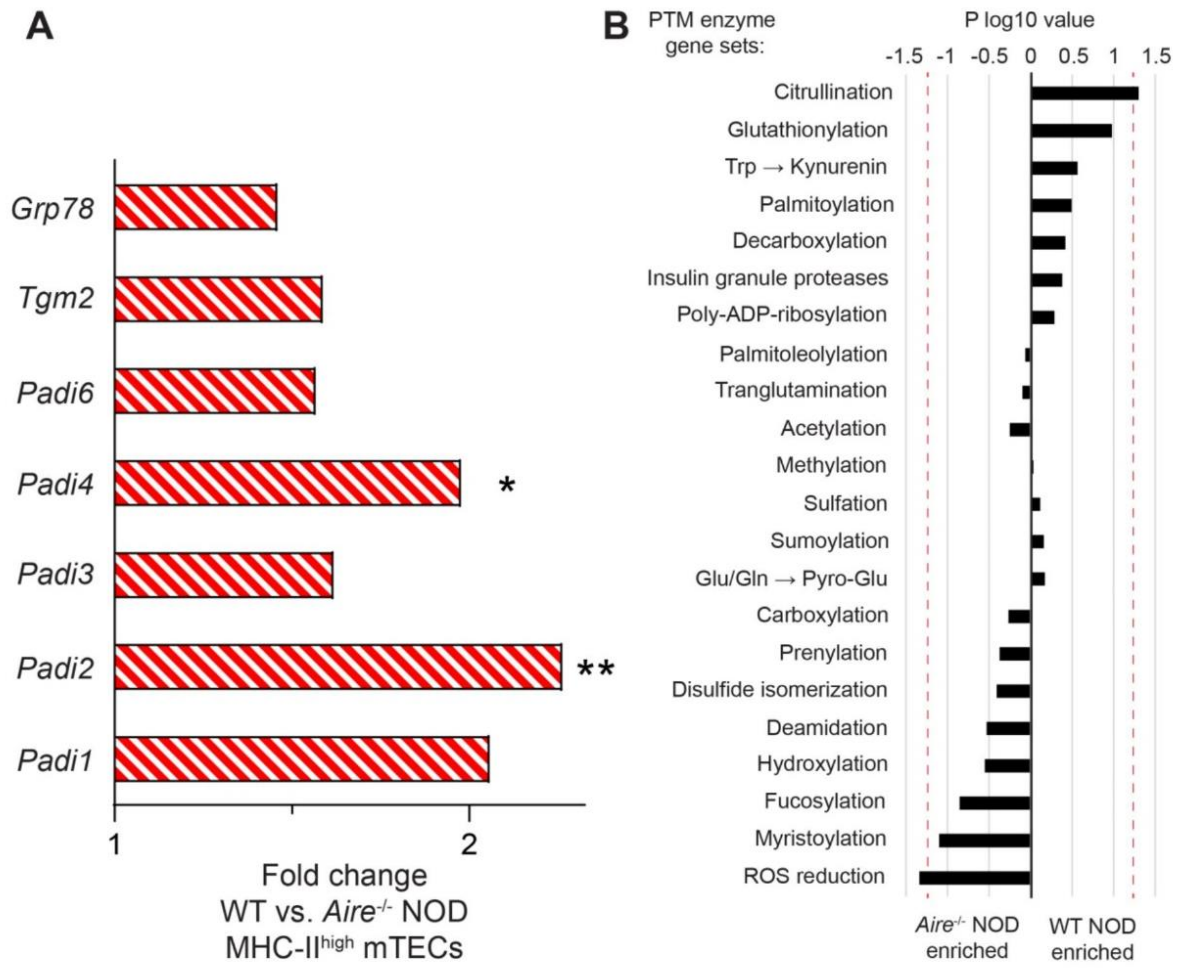
B



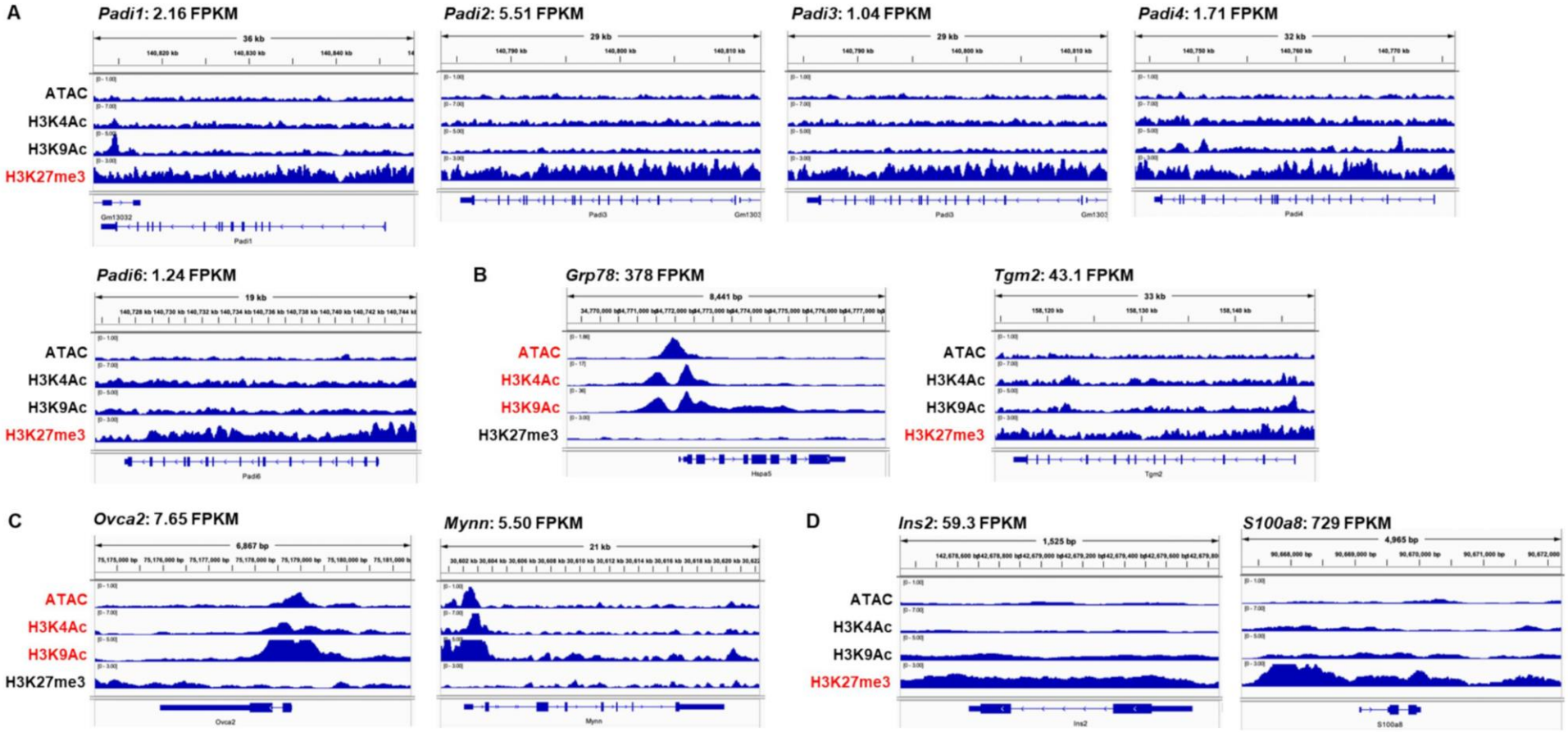
Supplementary Figure 4. Citrullinated peptides identified in mTECs of 10-week-old C57BL/6 mice. (A) List of the citrullinated peptides identified. **(B)** Spectrum of the R5-citrullinated myosin-9 peptide KQLVRQVR.



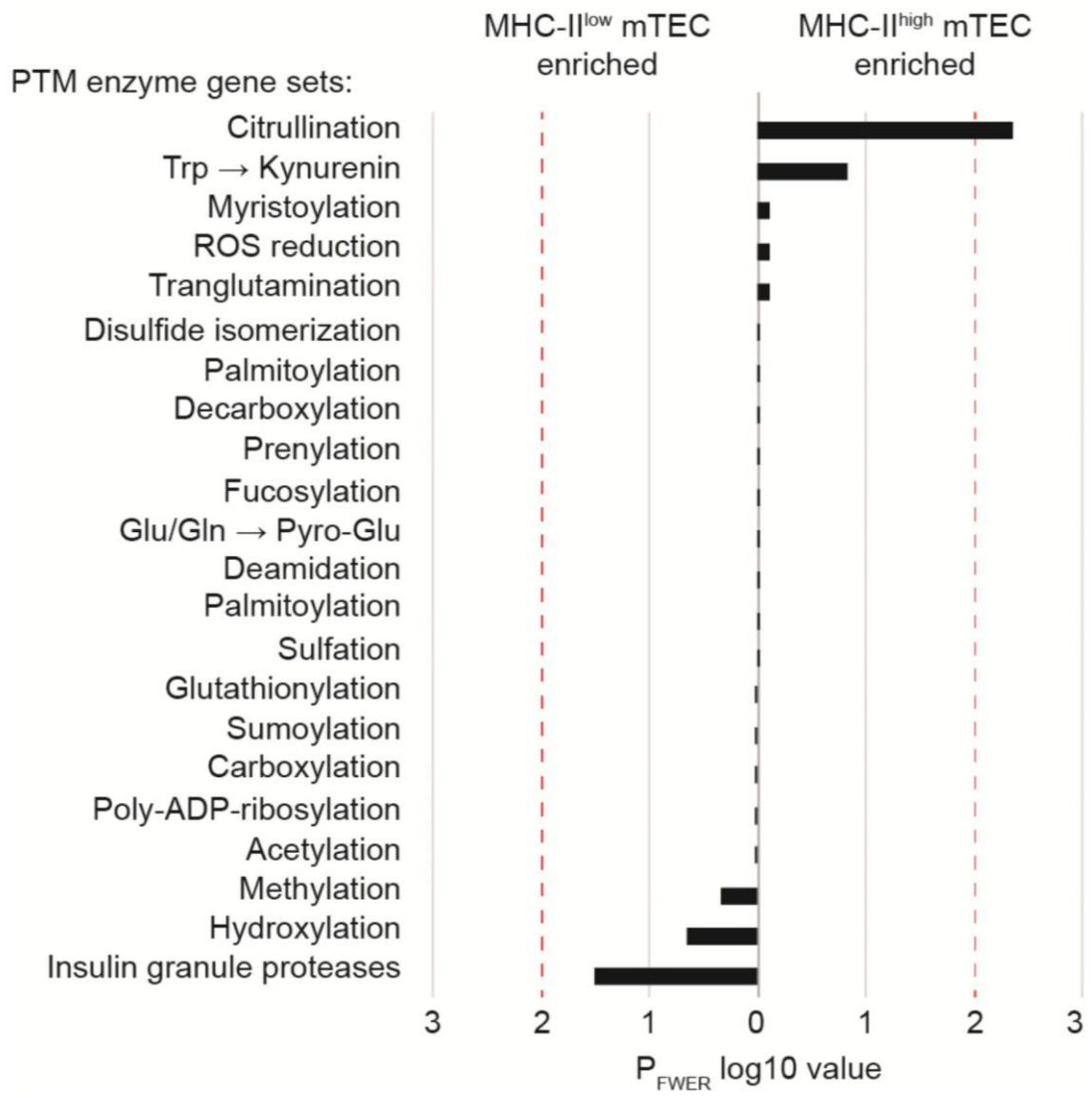
Supplementary Figure 5. Gene expression of PADI citrullinating enzymes in human immature HLA Class II (HLA-II)^{low} and mature HLA-II^{high} mTECs (n=3 donors). Data was extracted from our published RNAseq datasets (12).



Supplementary Figure 6. Enrichment of *Padi* isoform genes and PTM enzyme gene sets in mature MHC-II^{high} mTECs from WT vs. *Aire*$^{-/-}$ NOD mice (n=3/each). (A) Fold change expression of *Padi* isoform genes in MHC-II^{high} mTECs from 4-week-old WT vs. *Aire*$^{-/-}$ NOD mice analyzed by DNA microarray. **p*=0.04, ***p*=0.01 by false discovery rate method (GEO2R). **(B)** Enrichment of PTM enzyme gene sets in MHC-II^{high} mTECs from WT vs *Aire*$^{-/-}$ NOD mice. Red lines indicate the 0.05 *p* cut-off value for statistical significance [for citrullination, *p*=0.05, corresponding to a WT normalized enrichment score (NES) of 1.5]. Data is from GEO GSE12073 (8).



Supplementary Figure 7. ATAC-Seq and Chip-seq profiles of *Padi* genes in MHC-II^{high} mTECs from C57BL/6 mice. ATAC-seq and Chip-seq datasets (GSE114713) from Handel et al. (13) were analyzed. **(A)** *Padi* genes display three hallmarks of Aire-dependent expression: a closed chromatin ATAC profile (first row), near-absence of the activating chromatin modifications H3K4Ac and H3K9Ac (second and third row) and presence of the repressive modification H3K27me3 throughout the gene locus (fourth raw). **(B)** Conversely, the *Grp78* gene displays an open chromatin ATAC profile, presence of the activating chromatin modifications H3K4Ac and H3K9Ac at transcriptional start sites and absence of the repressive modification H3K27me3, in line with an Aire-independent expression. The *Tgm2* gene displays an intermediate profile. **(C-D)** Profiles of reference Aire-independent (C) and Aire-dependent genes (D) are shown as controls.



Supplementary Figure 8. Enrichment of PTM enzyme gene sets in immature MHC-II^{low} *vs* mature MHC-II^{high} mTECs from WT C57BL/6 mice (n=3). Red lines indicate the 0.01 P_{FWER} cut-off value for statistical significance (for citrullination, P_{FWER}=0.0045, corresponding to a MHC-II^{high} mTEC normalized enrichment score of 2.3).

	PE	PE-CF594	APC	BV650	BV711	BV786
PE		261X	PPI6-14	9R	9X	261R
PE-CF594			298R	298X	360R	360X
APC				435R	435X	502R
BV650					152	502X
BV711						Flu MP58-66
BV786						

Supplementary Table 1. Combinatorial HLA-A2 MMr panel used in Figure 2. See Supplementary Figures 1-2 for details.

	nPOD case	RRID	Sex	Age (yrs)	T1D T2D (yrs)	Positive aAbs	C-peptide (ng/ml)	Pancreas MMr+ cells						
								GRP78 9R	GRP78 9X	GRP78 360R	GRP78 360X	ZnT8 186-194	UCN3 1-9	MelA 26-35
T1D (n=10)	6070*	SAMN15879127	F	23	7	IA-2/mIAA	<0.05	0	16		53	74	39	0
	6113*	SAMN15879170	F	13	2	mIAA	<0.05	0	0		0			
	6161*	SAMN15879217	F	19	7	IA-2/mIAA	<0.05	12	45			124		0
	6211*	SAMN15879267	F	24	4	GAD/IA-2/ZnT8/mIAA	<0.05	3	4		0	30		0
	6212	SAMN15879268	M	20	5	mIAA	<0.05	0	0		0	0		
	6237	SAMN15879293	F	18	12	GAD/mIAA	<0.05	7	0		0	267	8	0
	6242	SAMN15879298	M	39	19	IA-2/mIAA	<0.05	0	0	0		66		0
	6258*	SAMN15879313	F	39	37	mIAA	<0.05	0	4	0		118	19	0
	6325*	SAMN15879379	F	20	6	GAD/IA-2	0.14	13	19	0		28	28	0
	060217	NA	F	39	21	GAD	NA	0	41	0			142	
aAb+ (n=9)	6080	SAMN15879137	F	69	NA	GAD/mIAA	1.84	0	0		0	55		25
	6101	SAMN15879158	M	65	NA	GAD	26.18	0	0		0	0		
	6123	SAMN15879180	F	23	NA	GAD	2.01	0			0	0		
	6151	SAMN15879207	M	30	NA	GAD	5.49	0	5		0	28	9	0
	6154	SAMN15879210	F	49	NA	GAD	<0.05	0	0	0		64	21	0
	6171	SAMN15879227	F	4	NA	GAD	8.95	0	7		8	37		0
	6347	SAMN15879401	M	9	NA	mIAA	3.26	6	49	7		33		0
	6388	SAMN15879441	F	25	NA	GAD/Miaa	1.38	0	7	0		34	0	0
	6397	SAMN15879450	F	21	NA	GAD	12.8	4	4	0		42	6	0
No diabetes (n=11)	6103	SAMN15879160	M	2	NA	—	0.98	30	0			55	0	0
	6174	SAMN15879230	M	21	NA	—	3.00	4	0		4			
	6179	SAMN15879235	F	20	NA	—	2.74	0	7		0	96	0	0
	6182	SAMN15879238	M	3	NA	—	2.28	0	4		0	23	4	0
	6227	SAMN15879283	F	17	NA	—	2.75	5	12		0	3	18	0
	6234	SAMN15879290	F	20	NA	—	6.89	0	0	0		6	0	0
	6271	SAMN15879325	M	17	NA	—	11.47	0	0	0		0		
	6287	SAMN15879341	F	57	NA	—	4.75	0	0	0		4		0
	6289	SAMN15879343	M	19	NA	—	8.05	0	0	0		0		
	6357	SAMN15879410	M	5	NA	—	8.82		128			0		
	6366	SAMN15879419	F	21	NA	—	0.41	0	35		0	0		
T2D/pancreatitis (n (n=4+1)	6028	SAMN15879085	M	33	17	—	22.40		0			0		
	6059	SAMN15879116	F	19	0.3	—	10.68	0	0			0		
	6273	SAMN15879327	F	45	2	—	3.17	3	7					
	6275	SAMN15879329	M	48	2	—	3.46	0	0			0		
	6439	SAMN15879492	M	27	NA	—	5.33	5	22					

Supplementary Table 2. nPOD cases analyzed by *in-situ* tissue MMr staining. The clinical characteristics of each case are reported along with the counts ($x10^{-3}$) of MMr^{+} cells/mm^2 pancreas section area for each of the indicated peptides. Positive sections are marked in red. Case #6287 (presenting a circumscribed neuroendocrine tumor in the pancreatic pan-body region; pan-tail region analyzed here) was classified as a non-diabetic control. T1D case #060217 is from the EUnPOD biobank. Control duodenum sections were also tested for T1D cases marked with an asterisk. NA, not applicable or not available; mIAA, micro-insulin aAbs. Results for epitopes other than GRP78 are from (12, 14).

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Acetylation	Kat2a	Q9JHD2	Q92830
	Naa10	Q9QY36	P41227
	Naa11	Q3UX61	Q9BSU3
	Naa15	Q80UM3	Q9BXJ9
	Naa16	Q9DBB4	Q6N069
	Naa20	P61600	P61599
	Naa25	Q8BWZ3	Q14CX7
	Naa30	Q8CES0	Q147X3
	Naa35	Q6PHQ8	Q5VZE5
	Naa38	Q9D2U5	Q9BRA0
	Naa50	Q6PGB6	Q9GZZ1
	Naa60	Q9DBU2	Q9H7X0
	Nat8	Q9JIY7	Q9UHE5
	Nat9	Q3UG98	Q9BTE0
	Sirt1	Q923E4	Q96EB6
	Sirt2	Q8VDQ8	Q8IXJ6
	Sirt7	Q8BKJ9	Q9NRC8
Carboxylation	Ggcx	Q9QYC7	A7YA96
Cisteinylation (non-enzymatic)	n/a		
Citrullination	Padi1	Q9Z185	Q9ULC6
	Padi2	Q08642	Q9Y2J8
	Padi3	Q9Z184	Q9ULW8
	Padi4	Q9Z183	Q9UM07
	Padi6	Q8K3V4	Q6TGC4
Deamidation	Ntan1	Q64311	Q96AB6
	Ntaq1	Q80WB5	Q96HA8
Decarboxylation	Gad1	P48318	Q99259
	Gad2	P48320	Q05329
	Gadl1	Q80WP8	Q6ZQY3
Fucosylation	Fut1	O09160	P19526
	Fut10 (Sec1)	Q5F2L2	Q6P4F1
	Fut11	Q8BHC9	Q495W5
	Fut2 (Sec2)	Q9JL27	Q10981
	Fut3	n/a	P21217
	Fut4	Q11127	P22083
	Fut7	Q11131	Q11130
	Fut8	Q9WTS2	Q9BYC5
	Fut9	O88819	Q9Y231
	Pofut1	Q91ZW2	Q9H488
	Pofut2	Q8VHI3	Q9Y2G5
Glutathionylation	Ggt1*	Q60928	P19440
	Ggt3p*	n/a	A6NGU5
	Ggt5*	Q9Z2A9	P36269
	Ggt6	Q6PDE7	Q6P531
	Ggt7	Q99JP7	Q9UJ14
	Gsta1	P13745	P08263
	Gsta2	P10648	P09210
	Gsta3	P30115	Q16772
	Gsta4	P24472	O15217
	Gsta5	n/a	Q7RTV2
	Gstk1	Q9DCM2	Q9Y2Q3
	Gstm1	P10649	P09488
	Gstm2	P15626	P28161
	Gstm3	P19639	P21266
	Gstm4	Q8R5I6	Q03013
	Gstm5	P48774	P46439
	Gstm6	O35660	n/a
	Gstm7	Q80W21	n/a
	Gsto1	O09131	P78417
	Gsto2	Q8K2Q2	Q9H4Y5
	Gstp1	P19157	P09211
	Gstp2	P46425	n/a
	Gstp3	Q8VC73	n/a
	Gstt1	Q64471	P30711
Glutathionylation	Gstt2	Q61133	P0CG29
	Gstt2b	n/a	P0CG30
	Gstt3	Q99L20	n/a
	Gstt4	Q9D4P7	A0A1W2PR19
	Lancl1	O89112	O43813
	Mgst1	Q91VS7	P10620
	Mgst2	A2RST1	Q99735
	Mgst3	Q9CPU4	O14880
Hydroxylation	Asph	Q8BSY0	Q12797
	Egln1	Q91YE3	Q9GZT9
	Egln2	Q91YE2	Q96KS0
	Egln3	Q91UZ4	Q9H6Z9
	Jmjd6	Q9ERI5	Q6NYC1
	Kdm8	Q9CXT6	Q8N371
Malonylation (non-enzymatic)	n/a		
Methylation	Antkmt	Q501J2	Q9BQD7
	Atpsckmt	Q9D1Z3	Q6P4H8
	Camkmt	Q3U2J5	Q7Z624
	Ehmt2	Q9Z148	Q96KQ7
	Icmt	Q9EQK7	O60725
	Kmt5a	Q2YDW7	Q9NQR1
	Lcmt1	A2RTH5	Q9UIC8
	Lcmt2	Q8BYR1	O60294
	Mettl11b	B2RXM4	Q5VVY1
	Mettl18	Q9CZ09	O95568
	Mettl21a	Q9CQL0	Q8WXB1
	Mettl21c	Q8BLU2	Q5VZV1
	Mettl21e	Q8CDZ2	n/a
	Mettl22	Q8R1C6	Q9BUU2
	N6amt1	Q6SKR2	Q9Y5N5
	Ndufaf7	Q9CWG8	Q7L592
	Ntmt1 (Mettl11a)	Q8R2U4	Q9BV86
	Pcmt1*	P23506	P22061
	Pcmtd1*	P59913	Q96MG8
	Prmt1	Q9JIF0	Q99873
	Prmt2	Q9R144	P55345
	Prmt3	Q922H1	O60678
	Prmt4 (Carm1)	Q9WVG6	Q86X55
	Prmt5 (Anm5)	Q8CIG8	O14744
	Prmt6 (Anm6)	Q6NZB1	Q96LA8
	Prmt7 (Anm7)	Q922X9	Q9NVM4
	Prmt8 (Anm8)	Q6PAK3	Q9NR22
	Prmt9 (Anm9)	Q3U3W5	Q6P2P2
	Setd2	E9Q5F9	Q9BYW2
	Setd4	P58467	Q9NVD3
	Setd7	Q8VHL1	Q8WTS6

continued

Supplementary Table 3. Mouse genes coding for enzymes involved in PTMs. This list was used for the analysis of gene sets presented in Fig. 5G and Supplemental Fig. 5. PTMs involving very large sets of enzymes were excluded, namely glycosylation, phosphorylation, oxidation and ubiquitination. This list does not include enzymes described to selectively modify specific proteins or histones. Asterisks indicate enzymes indirectly involved in PTMs.

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Myristoylation	Nmt1	O70310	P30419
	Nmt2	O70311	O60551
Nitration (non-enzymatic)	n/a		
Nitrosylation (non-enzymatic)	n/a		
Palmitoylation	Zdhhc1	Q8R0N9	Q8WTX9
	Zdhhc2	P59267	Q9UIJ5
	Zdhhc3	Q8R173	Q9NYG2
	Zdhhc4	Q9D6H5	Q9NPG8
	Zdhhc5	Q8VDZ4	Q9C0B5
	Zdhhc6	Q9CPV7	Q9H6R6
	Zdhhc7	Q91WU6	Q9NXF8
	Zdhhc8	Q5Y5T5	Q9ULC8
	Zdhhc9	P59268	Q9Y397
	Zdhhc11	Q14AK4	Q9H8X9
	Zdhhc11b	n/a	P0C7U3
	Zdhhc12	Q8VC90	Q96GR4
	Zdhhc13	Q9CWU2	Q8IUH4
	Zdhhc14	Q8BQQ1	Q8IZN3
	Zdhhc15	Q8BGJ0	Q96MV8
	Zdhhc16	Q969W1	Q969W1
	Zdhhc17	Q80TN5	Q8IUH5
	Zdhhc18	Q5Y5T2	Q9NUE0
	Zdhhc19	Q810M5	Q8WVZ1
	Zdhhc20	Q5Y5T1	Q5W0Z9
	Zdhhc21	Q9D270	Q8IVQ6
	Zdhhc22	A0PK84	Q8N966
	Zdhhc23	Q5Y5T3	Q8IYP9
	Zdhhc24	Q6IR37	Q6UX98
Palmitoleoylation	Porcn	Q9JJJ7	Q9H237
	Notum	Q8R116	Q6P988
Poly-ADP-ribosylation	Parp1	P11103	P09874
	Parp2	O88554	Q9UGN5
	Parp3	Q3ULW8	Q9Y6F1
	Parp4	E9PYK3	Q9UKK3
	Parp6	Q6P6P7	Q2NL67
	Parp8	Q3UD82	Q8N3A8
	Parp9	Q8CAS9	Q8IXQ6
	Parp10	Q8CIE4	Q53GL7
	Parp11	Q8CFF0	Q9NR21
	Parp12	Q8BZ20	Q9H0J9
	Parp14	Q2EMV9	Q460N5
	PARP15	n/a	Q460N3
	Parp16	Q7TMM8	Q8N5Y8
	Tiparp (Parp7)	Q8C1B2	Q7Z3E1
	Tnks (Parp5a)	Q6PFX9	O95271
	Tnks2 (Parp5b)	Q3UES3	Q9H2K2
Prenylation	Fnta	Q61239	P49354
	Fntb	Q8K2I1	P49356
	Pggt1b	A0A494BAX1	P53609
	Rce1	P57791	Q9Y256

PTM	Mouse gene	Mouse UniProt #	Human Uniprot #
Glu/Gln → Pyro-Glu	Qpct	Q9CYK2	Q16769
	Qpctl	Q8BH73	Q9NXS2
Disulfide isomerization	Creld1	Q91XD7	Q96HD1
	Creld2	Q9CYA0	Q6UXH1
	Ero1a*	Q8R180	Q96HE7
	Ero1b*	Q8R2E9	Q86YB8
	P4hb (Pdia1)	P09103	P07237
	Pdia2	D3Z6P0	Q13087
	Pdia3	P27773	P30101
	Pdia4	P08003	P13667
	Pdia5	Q921X9	Q14554
	Pdia6	Q922R8	Q15084
	Qsox1	Q8BND5	O00391
	Qsox2	Q3TMX7	Q6ZRP7
	Tmx3	Q8BXZ1	Q96JJ7
ROS reduction	Cat (Cata)*	P24270	P04040
	Cybb*	Q61093	P04839
	Mpo*	P11247	P05164
	Nos1*	Q9Z0J4	C9J5P6
	Nos2*	P29477	P35228
	Sod1*	P08228	P00441
	Sod2*	P09671	P04179
	Sod3*	O09164	P08924
Sulfation	Tpst1	O70281	O60507
	Tpst2	O88856	O60704
Sulfone (non-enzymatic)	n/a		
Sumoylation	Sae1	Q9R1T2	Q9UBE0
	Uba2 (Sae2)	Q9Z1F9	Q9UBT2
Transglutamination	Tgm1	Q9JLF6	P22735
	Tgm2	P21981	P21980
	Tgm3	Q08189	Q08188
	Tgm4	Q8BZH1	P49221
	Tgm5	Q9D7I9	O43548
	Tgm6 (Tgm3l)	Q8BM11	O95932
	Tgm7	A2ART8	Q96PF1
Trp → Kynurenine	Afmid	Q8K4H1	Q63HM1
	Ido1	P28776	P14902
	Ido2	Q8R0V5	Q6ZQW0
	Tdo2 (T23o)	P48776	P48775
Insulin granule proteases	Cpe	Q00493	P16870
	Cpn1	Q9JJN5	P15169
	Cpn2	Q9DBB9	P22792
	Ctsb	P10605	P07858
	Ctsd	P18242	P07339
	Ctsf	Q9R013	Q9UBX1
	Ctsl	P06797	P07711
	Dpp7	Q9ET22	Q9UHL4
	Furin	P23188	P09958
	Pcsk1	P63239	P29120
	Pcsk2	P21661	P16519
	Tpp1	O89023	O14773

SUPPLEMENTARY REFERENCES

1. Buitinga M, Callebaut A, Marques Camara Sodre F, Crevecoeur I, Blahnik-Fagan G, Yang ML, Bugliani M, Arribas-Layton D, Marre M, Cook DP, Waelkens E, Mallone R, Piganelli JD, Marchetti P, Mamula MJ, Derua R, James EA, Mathieu C, Overbergh L: Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. Diabetes 2018;67:2337-2348

2. Callebaut A, Derua R, Vig S, Delong T, Mathieu C, Overbergh L: Identification of Deamidated Peptides in Cytokine-Exposed MIN6 Cells through LC-MS/MS Using a Shortened Digestion Time and Inspection of MS2 Spectra. J Proteome Res 2021;20:1405-1414

3. Guyon C, Jmari N, Padonou F, Li YC, Ucar O, Fujikado N, Coulpier F, Blanchet C, Root DE, Giraud M: Aire-dependent genes undergo Clp1-mediated 3'UTR shortening associated with higher transcript stability in the thymus. eLife 2020;9

4. Langmead B, Trapnell C, Pop M, Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 2009;10:R25

5. Quinlan AR, Hall IM: BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 2010;26:841-842

6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102:15545-15550

7. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol 2010;11:R106

8. Guerau-de-Arellano M, Martinic M, Benoist C, Mathis D: Neonatal tolerance revisited: a perinatal window for Aire control of autoimmunity. J Exp Med 2009;206:1245-1252

9. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP: GenePattern 2.0. Nat Genet 2006;38:500-501

10. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003;4:249-264

11. Serge A, Bailly AL, Aurrand-Lions M, Imhof BA, Irla M: For3D: Full organ reconstruction in 3D, an automatized tool for deciphering the complexity of lymphoid organs. J Immunol Methods 2015;424:32-42

12. Gonzalez-Duque S, Azoury ME, Colli ML, Afonso G, Turatsinze JV, Nigi L, Lalanne AI, Sebastiani G, Carre A, Pinto S, Culina S, Corcos N, Bugliani M, Marchetti P, Armanet M, Diedisheim M, Kyewski B, Steinmetz LM, Buus S, You S, Dubois-Laforgue D, Larger E, Beressi JP, Bruno G, Dotta F, Scharfmann R, Eizirik DL, Verdier Y, Vinh J, Mallone R: Conventional and Neo-antigenic Peptides Presented by beta Cells Are Targeted by Circulating Naive CD8+ T Cells in Type 1 Diabetic and Healthy Donors. Cell Metab 2018;28:946-960 e946

13. Handel AE, Shikama-Dorn N, Zhanybekova S, Maio S, Graedel AN, Zuklys S, Ponting CP, Hollander GA: Comprehensively Profiling the Chromatin Architecture of Tissue Restricted Antigen Expression in Thymic Epithelial Cells Over Development. Front Immunol 2018;9:2120

14. Culina S, Lalanne AI, Afonso G, Cerosaletti K, Pinto S, Sebastiani G, Kuranda K, Nigi L, Eugster A, Osterbye T, Maugein A, McLaren JE, Ladell K, Larger E, Beressi JP, Lissina A, Appay V, Davidson HW, Buus S, Price DA, Kuhn M, Bonifacio E, Battaglia M, Caillat-Zucman S, Dotta F, Scharfmann R, Kyewski B, Mallone R, ImMaDiab Study G: Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. Sci Immunol 2018;3:eaao4013

Annex 5: Original article on the benchmarking of a novel MARS algorithm to identify *cis*-spliced peptides

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

MARS: Improved De Novo Peptide Candidate Selection for Non-Canonical Antigen Target Discovery in Cancer

Hanqing Liao
University of Oxford
Carolina Barra
Technical University of Denmark https://orcid.org/0000-0002-6836-4906
Xu Peng
University of Oxford
Isaac Woodhouse
University of Oxford https://orcid.org/0000-0002-8221-7697
Arun Tailor
University of Oxford
Robert Parker
University of Oxford
Alexia Carré
Institute Cochin, INSERM U1016, CNRS UMR8104 https://orcid.org/0000-0001-8144-7194
Roberto Mallone
Institute Cochin, INSERM U1016, CNRS UMR8104 https://orcid.org/0000-0002-9846-8861
Michael Hogan
Children's Hospital of Philadelphia https://orcid.org/0000-0002-6244-2993
Wayne Paes
University of Oxford https://orcid.org/0000-0002-0529-2765
Laurence Eisenlohr
University of Pennsylvania https://orcid.org/0000-0002-8475-7910
Morten Nielsen
Technical University of Denmark https://orcid.org/0000-0001-7885-4311
Nicola Ternette (✉ nicola.ternette@ndm.ox.ac.uk)
University of Oxford https://orcid.org/0000-0002-9283-0743

Article

Keywords:

Posted Date: August 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1890352/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

MARS: Improved *De Novo* Peptide Candidate Selection for Non-Canonical Antigen Target Discovery in Cancer

Hanqing Liao[1,2], Carolina Barra[3], Xu Peng[1], Isaac Woodhouse[1,2], Arun Tailor[1,2], Robert Parker[1,2], Alexia Carré[4], Roberto Mallone[4], Persephone Borrow[2], Michael J Hogan[5], Wayne Paes[1,2], Laurence C. Eisenlohr[5,6], Morten Nielsen[3]*, Nicola Ternette[1,2*#]

[1] The Jenner Institute, University of Oxford, OX37BN, UK

[2] Centre for Immuno-Oncology, Nuffield Department of Medicine, University of Oxford, OX37DQ, UK

[3] Technical University Denmark, Copenhagen, Denmark

[4] Université de Paris Cité, Institut Cochin, CNRS, INSERM, 75014 Paris, France

[5]Department of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, 19104.

[6]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104.

*equal contribution

to whom correspondence may be addressed: nicola.ternette@ndm.ox.ac.uk

ABSTRACT

Cryptic Human Leukocyte Antigen (HLA)-presented peptide identification from unannotated genome sources is a priority for target antigen discovery for development of next generation immunotherapies in cancer. Current immunopeptidomic approaches utilize the integration of transcriptomics data to inform spectral interpretation, however, recent observations that tumour-associated antigen-encoding RNA levels are often low highlights limitations of such proteogenomic approaches[1].

We here employ a *de novo* sequencing approach with a refined, MHC-centric analysis strategy to detect non-canonical HLA-associated peptide sequences (HLAp) in cancer without integration of transcript sequence information. Our strategy integrates HLA binding prediction, peptide retention time prediction, and average local confidence scores culminating in the machine learning model MARS (**M**HC binding prediction, **A**verage Local Confidence Score, and **R**etention time integration for improved *de novo* candidate **S**election).

We demonstrate increased HLA-I peptide identification sensitivity by benchmarking our model against *de novo* sequencing alone with a large synthetic HLA-I peptide library dataset. We further define the sensitivity of MARS by reanalysis of a published dataset of high-quality non-canonical HLAp identifications in human cancer cell line and tissue datasets and achieve almost 2-fold improvement of the full sequence recall (FSR) for high quality spectral assignments in comparison to *de novo* sequencing alone[2]. We minimize the false discovery rate (FDR) through a step-wise peptide sequence mapping strategy and are able to expand the reported non-canonical peptide space with an assignment accuracy above 85.7%.

Finally, we utilize MARS to detect and validate lncRNA-derived peptides in human cervical tumour resections, demonstrating its suitability to discover novel, non-canonical peptide sequences in primary tumour tissue at reduced FDR, in the absence of transcriptomic sequencing data.

INTRODUCTION

The analysis of HLA-bound peptide ligands using liquid chromatography mass spectrometry (LC-MS) technology (immunopeptidomics) is a rapidly evolving field that has advanced identification of T cell antigens including clinically relevant epitopes in autoimmunity, pathogen infection and cancer [4-6]. These approaches have recently been expanded to identification of peptides from unannotated open reading frames (ORFs), or out-of-frame translated sequences, that are valuable for development of immunotherapeutic approaches to cancer, using the integration of RNA sequence information [2, 4, 7-9]. However, recent observations that many HLA-presented peptides exhibit minimal RNA expression highlight the need for transcriptomics data-independent interrogation of the immunopeptidome for cancer antigen discovery.

Software tools to interrogate LC-MS spectral data have widely been adopted from the field of proteomics, and extensive differences in performance have been observed for HLA-derived peptide data [2, 10, 11]. Spectra interpretation software is based on three main strategies. The most common approach, the "database search" approach [12-14], relies on matching the masses of experimentally obtained peptide fragment spectra to all theoretically possible peptide fragments of sequences in a protein sequence list or "database". This approach is limited to identifications in the sequence database, and can therefore not be used for discovery of non-canonical peptides unless such information is added to known canonical protein databases. For increased sensitivity, this approach has been extended by integration of spectral matching assignment, which uses advanced deep learning and peptide fragmentation prediction to annotate the experimentally obtained peptide fragment spectra with high accuracy [15-19]. Thirdly, *de novo* algorithms interpret the MS spectra without comparison to a protein sequence database, annotating the amino acid sequence directly from the spectral information available [20-22]. Integration of *de novo* approaches and database searching (*de novo*-assisted database approach) has been shown to be particularly advantageous for HLA-peptide datasets [10, 11, 23].

For the interrogation of immunopeptidome datasets with a focus on cancer-specific peptide ligands, proteogenomics approaches have been used to refine the database-search approach in order to identify peptide sequences that contain tumour-specific mutations or other non-canonical sources like functional RNA genes, intronic sequences, and/or reading frames in unannotated genomic origins that are generally not present in universal protein databases [2, 7, 23-25]. Here, genomics and/or transcriptomics data is integrated to achieve a highly accurate, sample-specific database that can be used as template for a "database search" analysis approach.

The incorporation of "personalised" binding prediction (MSRescue [26]), or difference between observed and predicted retention time (DeepRTplus [27], Deep Rescore [19]) to database-search based peptide identification workflows can also improve performance. Recently, machine learning has been applied in a personalized fashion in order to identify patient neoantigens [28].

The purpose of this study is to increase the power of the *de novo* approach for identification of non-canonical HLA-I-associated peptides (HLA-Ip) without the use of a personalized template protein database. We report that by integration of (i) HLA binding affinity predictions and (ii) differences between observed and predicted retention time (RT) of candidate peptide sequences as additional factors, HLA-I peptide identification sensitivity can be improved over the baseline *de novo* sequencing annotation. We validate our findings by benchmarking our algorithm termed MARS on LC-MS data from synthetic HLA-I peptide libraries, and published non-canonical peptide sequences that were obtained from proteogenomic analyses. We then proceed and develop a hierarchical strategy to map MARS sequence candidates to their likely origin in the genome, and conclude in a final false discovery rate below 14.3% - offering opportunity for improved shortlisting of accurate sequence identifications from the generally observed global 65% FDR in *de novo* sequence data[3]. Through this approach, we are able to expand the non-canonical peptidome of published data by hundreds of peptides. We finally detect and validate

lncRNA derived peptides in primary human tumour tissue, demonstrating applicability for antigen discovery in cancer.

RESULTS

Training Data Selection and Processing

Our aim was to construct and validate our computational model for improved *de novo* (template-independent) identification of HLA-I-associated peptides. We selected a panel of in-house single allele (CD4.221 cells expressing either A*01:01, A*02:01, A*03:01, A*11:01, B*08:01, B*44:02, B*57:03, C*03:03, or C*03:04) and multi allele (Jurkat cell line: A*03:01, B*07:02, B35:03, C*04:01, C*07:02; and the C1866 cell line: A*24:02. B*40:02, B*51:01, C*03:03, C*14:02) data sets to develop models and to evaluate model performances.

We analyzed these datasets using Peaks *DeNovo* and also the Peaks *DeNovo*-assisted Database search (DB) using the SwissProt human canonical proteome [29]. From the Peaks *DeNovo* search, we obtained up to 100 sequence candidates per spectrum (independent from the provided sequence database), with associated average local confidence (ALC) scores for each candidate sequence as a measure for the confidence of each sequence assignment. The Peaks DB search results were returned as a list of sequence assignments that matched to the SwissProt database. This latter search results in one sequence assignment per spectrum (with exception of L/I isomers) with an associated score (-10*log*P), which reflects the probability of a correct peptide spectrum match. The results are controlled by a simultaneous decoy database search which provides an accurate estimation of the false discovery rate to be expected in the result [30].

In order to define a subset of "true sequences" to use as a training dataset, we defined spectra as confidently assigned if they were identified with Peaks DB with high confidence -10*log*P ≥ 20 resulting in an FDR of less than 1% on peptide-spectrum match (PSM) level. In addition, we only included peptides

with a length of 8-13 amino acids, and those that were assigned a low NetMHCpan 4.1 Eluted Ligand Rank (ELR) score (*ELR* < 10), aiming to remove peptides that are likely originating from co-purification during biochemical enrichment. All spectra that fulfilled these criteria were defined as the training dataset. Next, within the Peaks *DeNovo* search results, the "true" peptide was considered the Peaks DB assigned sequence identification with all constituent isoleucine (I) replaced by leucine (L), because Peaks *DeNovo* algorithm does not distinguish I from L. Only those spectra with the "true" (DB) peptides present in the list of the *de novo* peptide candidates were selected for model development and evaluation. In order to confirm that the I to L replacement does not negatively affect the NetMHCpan binding prediction, we tested NetMHCpan predictions within the test set for both L and I variants, and found that replacement of all I with L incurred significantly less divergence from the original result than *vice versa* (Supplementary Figure 1).

All selected spectra were then split into 5 equal parts to form a five-fold cross-validation scheme, such that no two parts shared any common 8-mer substring training peptide sequence. The model trainings and evaluations were carried out using a five-fold cross-validation scheme, where each time a different combination of four parts was used as model-training set, while the performances were evaluated on the fifth (test) partition. Finally, all five-fold test set predictions were concatenated to obtain the final model performance.

Two factor model evaluation: Incorporation of NetMHCpan prediction improves *de novo* candidate selection

We used the Peaks *DeNovo* result output as the baseline model. Accordingly, for each spectrum, the peptide candidate of the highest ALC was taken as the "identified" peptide further referred to as "ALC" result. In order to understand whether we could re-rank the *de novo* candidates by selecting the best candidate based on their binding prediction to the respective MHC allele of origin, we determined the

NetMHCpan binding percentile Rank score for each of the 100 sequence candidates in the PEAKS *DeNovo* list, and calculated for each peptide candidate an affine combination of the ALC and NetMHCpan predicted Rank (ELR) scores to produce a new composite score. Then, for each spectrum, the peptide candidate of the highest composite score was taken as the identified peptide. Because this affine combination was controlled by a single weight parameter, alpha, and modelled ALC and ELR as two peptide identity determining factors, we named it "two-factor alpha model". We calculated the F-ranks which describe the rank of the true peptide sequence in descending order, divided by the total number of candidates, for each spectrum as an identification performance measure. We applied a grid search to obtain the optimal alpha value for a minimal mean F-rank within the training set. Then this alpha value was applied to the test for performance evaluation. All five test sets yielded one-sided paired Wilcoxon signed-rank test p-values <2.2E10-16, indicating that the reduction of F-ranks by applying the two-factor alpha model in comparison to the ALC baseline model was highly significant (Supplementary Table 1, Figure 1A, B).

To examine whether the improved F-rank led to higher identification sensitivity when top scoring peptide candidate per spectrum was taken as its identity, we employed another measurement, the full sequence recall (FSR), which depicts the proportion of identified (top one scoring) peptides being correct. A paired T-test using the mean FSRs from each of the five partitions also suggested that two-factor alpha model (M=0.782, SD=0.00461) significantly improved FSR over the baseline ALC model (M=0.731, SD=0.00554); t(4)=36.74, p=3.28E10-6 (Figure 1C).

Lastly, the optimal alpha obtained in the five-fold cross-validations were highly concordant, suggesting the two-factor alpha model was robust (Supplementary Table 1). In summary, these results suggested that incorporating ELR as a factor can improve PSM candidate selection, resulting in improved HLAp peptide identification compared to ALC alone.

Incorporation of RT prediction further improves *de novo* candidate selection: Final model evaluation

To further improve PSM candidate selection, we incorporated DeepRTplus [27] to predict the RT of input peptides. We trained the DeepRTplus model with our training data, and applied the learned model to the training data itself to derive an absolute value of differences between predicted and observed retention time (DRT). We then defined an affine combination of the ALC, ELR and DRT scores to define a three-factor composite model and used the grid search to find optimal alpha (weight on ELR) and beta (weight on DRT) values to minimize the mean F-rank in the training data. Finally, the optimal values of alpha and beta were used to assign combinatory scores to each peptide in the test data.

We then proceeded with our evaluation strategy as before: a paired Wilcoxon test revealed that the F-ranks of the three-factor model were significantly lower than those of the two-factor alpha model (maximal of five p-values < 1.2 x 10E-10 (Supplementary Table 1, Figure 1B)), suggesting that the three-factor model outperformed the two-factor model in promoting the true peptide identity amongst all other candidates in each spectrum. From the perspective of FSR, the performance gain was consistent with the F-Rank: the FSR by the three-factor model (M=0.786, SD=0.00601) was higher than FSR of two-factor alpha model (M=0.781, SD=0.00462) t(4)=3.14, p=0.0349 (Figure 1C). Further, the model "training" was found to be robust in that the optimal alpha and beta were stable across test sets (Supplementary Table 1).

We concluded that adding in RT prediction as a factor can further improve PSM candidate selection, and fixed the alpha and beta of the model to 0.71 and 0.08, respectively. The resulting model was termed "MARS" as abbreviation for **M**HC-prediction, **A**LC and **R**etention-time prediction for the **S**election of *de novo* sequencing candidates.

MARS improves HLAp identification sensitivity across HLA-A and HLA-B alleles

To substantiate our findings, we applied MARS to synthetic HLAp standard data for sensitivity evaluation. We hypothesized that the improvement of the re-ranking of *de novo* HLAp identification by the MARS

model is robust in the sense that it can be applied to different datasets with fixed factor coefficients. The only requirement is a transfer learning scheme needed to calibrate the DeepRTplus model in order to accommodate the discrepancy introduced by the LC system setups and gradient conditions between different datasets.

We utilized a large dataset published by Wilhelm, Zolg et al., of ~169,000 synthetic HLA class I peptides acquired by liquid chromatography mass spectrometry (LC-MS) [17]. Out of this dataset, we selected the 116 most frequent HLA alleles [31], comprising 35 HLA-A, 59 HLA-B, and 22 HLA-C alleles (Supplementary Table 2). We here included spectra from the 3xHCD method only, which was reported the best performing in the original study. Each peptide's HLA binding information was extracted from IEDB [32, 33]. We further applied a filter to exclude peptide-HLA associations with NetMHCpan ELR>10, which we deemed likely inaccurate assignments. This resulted in a collection of 28,919 MS2 spectra from 24,193 unique peptide sequences, of which the majority of sequences (26,452 spectra/22,324 peptide sequences; 91.5%) were listed in the Peaks *DeNovo* result candidate list (hence forward referred to as "recoverable") (Supplementary Table 2). We selected sequence-RT pairs from peptide candidates with ALC > 95 to calibrate the DeepRTplus model, while keeping the alpha and beta values constant.

MARS outperformed ALC and the average F-ranks were significantly lower for MARS than ALC by paired Wilcoxon test ($p < 2.2 \times 10E-16$, Figure 2A). MARS FSR (M=0.907, SD=0.0344) was also significantly higher than ALC FSR (M=0.859, SD=0.0744): $t(115)=8.4295$, $p=1.147e-13$ (Figure 2B). MARS could again increase the number of identified peptide sequences over ALC (Figure 2C), and overall improvement was 7.23%. By grouping HLA alleles based on gene locus (A, B, and C), it can be seen that for HLA-A and B loci, the peptide FSRs can be significantly improved by MARS, but for HLA-C, the difference between ALC FSR and MARS FSR did not reach significance, possibly due to a lower number of spectra included for HLA-C (Figure 2D, Supplementary Table 2). When we further looked at the improvement of sequence identifications for all individual alleles, we noted that for HLA-A, only HLA-A*01:01 and HLA-A*31:01 did not benefit from

206 MARS over ALC, while MARS improved results for all other HLA-A assigned peptide sequences. A high
207 variability in performance across different HLA-B alleles was observed, with striking improvement in
208 peptide identifications using MARS for HLA-B*27:05 and other individual HLA-B alleles, while HLA-B35 did
209 not benefit from MARS rescoring (Figure 2D).
210
211 **MARS identifies previously identified HLAp of non-canonical origins without integration of sample-**
212 **specific RNA sequencing data**
213 To further assess the identification performance of the MARS model, and in particular the performance
214 in the non-canonical HLA peptide space, we utilized a published dataset from Chong *et al*. comprising nine
215 LC-MS datasets for five melanoma cell lines and two sets of lung cancer and healthy control tissues
216 spanning a wide range of HLA alleles [2]. The authors identify and report 508 non-canonical, HLA-presented
217 peptides mostly sourced from lncRNA genes, which are not present in the canonical human proteome.
218 The authors achieved identification of these peptides sequences using an immunopeptidomics approach
219 which carefully integrates RNA sequencing and ribosome profiling data for generation of personalized,
220 sample-specific transcript assemblies. These were then used for in-silico translation into protein
221 databases and mapping of non-canonical, HLA-presented sequences in the immunopeptidomics data
222 using a standard database-search.
223 Here, we wanted to understand whether MARS could recover these sequences without the integration of
224 any of the additional RNA sequencing datasets used in the original publication. Firstly, we searched the
225 Chong *et al.* datasets against the SwissProt human protein database using Peaks software in order to
226 exclude spectra that were assigned to the canonical human proteome from our search. N-terminal
227 acetylation and oxidation were included in the search parameters in order to match the analysis settings
228 used in the original publication. We identified a total of 83,295 canonical peptides using the PEAKS

database search then applied the MARS model to the remaining spectra, returning an additional 333,011 unique peptide identifications.

Overall, MARS was able to recover 227 of the 508 (45%) reported lncRNA derived peptides reported by Chong *et al.*. To note that only 272 out of the 508 peptides were listed as a candidate in the Peaks DeNovo search results, therefore the other sequences were "non-discoverable" by MARS. MARS therefore recovered 83% (227/272) of the possible "*de novo*-identifiable" space.

A clear separation of the number of peptides recovered with increasing false discovery rate was observed when applying different MARS score thresholds across all nine samples in comparison to ALC alone (Figure 3A). We therefore decided to define three major score confidence regions across this space, defining a MARS score threshold for increasing FDR: region 1: <30% FDR (MARS score >94.7), region2: >30% FDR (MARS score >89.7), and region 3: >45% FDR (MARS score >80) (Figure 3A).

With a stringent MARS score threshold of >94.7, and an estimated 30% FDR, we were able to recover 175/272 (64%) of sequences, while ALC alone could recover only 101/272 (37%) peptides (Figure 3B, Table S2, Supplementary Table 3), resembling an almost 2-fold improvement in FSR.

Integration of a sequence mapping strategy for MARS peptide candidates confirms non-canonical HLAp origin.

We next aimed to assign the 333,011 MARS identified sequences in the interrogated dataset to their respective likely gene origin. We therefore developed a stratified search approach to associate MARS identified HLAp to a wide range of possible origins across the human genome. We defined peptides mapping to (i) canonical origins mapping to known protein isoforms as "Human Protein", (ii) single amino acid substitutions as "1AA substitution", and (iii) other locations in the human genome as "non-canonical" (Figure 3C, and Methods section).

252 Further to the 83,295 peptides that had initially been identified by the PEAKS database search, 19,619
253 sequences were additionally identified as isoforms of the known proteome space as assigned and
254 reported by SwissProt, Tremble and Ensemble by MARS. A further 22,934 MARS sequences mapped to
255 canonical human proteins with exactly 1 aa substitution, which we considered as likely products of
256 germline or tumour-specific mutations (neoantigens). A total of 11,339 peptides were linearly mapped to
257 non-canonical sequences in the human genome. For 279,119 MARS peptides we could not find an
258 accurate match within the human genome with less than 1 aa distance, and we therefore considered
259 these sequences as distant from the human proteome ("unmatched", Figure 3D, left panel, see
260 Discussion).
261 193/272 of the Chong *et al.* published *de novo*-identifiable peptides were as expected part of the "non-
262 canonical" peptide category as defined here. In addition, with our search strategy, we found 16 peptides
263 reported in our DB database search, an additional 23 peptides matched to known human isoforms
264 reported by SwissProt, Trembl and/or Ensemble, and 30 peptides to human proteins with 1 AA
265 substitution (Figure 3D, right panel).
266 The 194 peptides identified by both Chong *et al.* and 194 MARS had an overall higher MARS score (mainly
267 falling into regions 1 and 2) in comparison to the 13,724 additional non-canonical peptide candidates
268 MARS recovered from the published datasets (Figure 3D, E), indicating that these peptides were recovered
269 with high confidence and that a higher MARS score cut-off can likely limit the FDR significantly.
270
271 **MARS extends the non-canonical peptide space by hundreds of peptides at an FDR <14.3%**
272 In order to determine the actual FDR experimentally for the novel MARS non-canonical peptide
273 identifications, we selected 97 peptides that closely resembled the observed MARS score distribution and
274 synthesized peptides for spectral validation and FDR evaluation (Figure 4A, Supplementary Figure 3). We
275 also included 6 peptides identified by both Chong *et al.* and MARS, and 3 peptides matching to the human

proteome as (likely positive) controls. We found 36 out of 97 sequences to be assigned correctly, resulting in an FDR of 14.3% (region 1, MARS score >94.8), 45.7% (region 2, MARS score > 89.7), and 66.0% (region 3, MARS score >80) (Figure 3F,G, Supplementary Table 6). All 9 positive control peptides were confirmed. When considering only peptides that were matched to a genome origin (excluding all "unmatched" identifications), there were 42 peptides selected, the FDR could further be reduced to 0 % (region 1, MARS score >94.8), 39.5% (region 2, MARS score > 89.7), and 63.0% (region 3, MARS score >80) (Figure 3G). This meant that all peptides with a MARS score of >94.8 that mapped to a genome origin were validated in this experiment.

We then proceeded to further stratify the non-canonical peptides using GENECODE "Biotype" definitions, including functional RNA genes and lncRNA genes, protein coding regions, unannotated regions, and pseudogenes as detailed in the Method section (Figure 3H). For this, we did not apply a hierarchical strategy that would consider certain origins more likely than others, but argue that without sample specific genome/RNA level evidence, peptides with multiple possible origins cannot be faithfully assigned to a single source.

Out of the 508 peptides originally published by Chong et al., 399 peptides are defined as "non-canonical" according to our categorization, which excludes peptides with single amino acid substitutions (57 peptides, "1 AA Substitutions" category) and those that map to canonical sequence databases in our workflow (52 peptides, "Human Protein" category). MARS was able to expand the originally published, in our definition "non-canonical" peptide identifications from 399 originally published peptides in this category to 1870 peptides from non-canonical origins at a an FDR between 0<FDR<14.3% (Figure 3H).

MARS uniquely identifies novel non-canonical lncRNA-derived HLA peptide sequences in primary tumour tissue

In order to establish whether MARS could identify novel non-canonical, lncRNA-derived peptides in primary tumour tissue, we applied MARS to a set of in-house immunopeptidomics data from ten primary cervical tumour resections spanning 35 HLA alleles (Supplementary Figure 2, Supplementary Table 5). Apart from 32,366 Peaks database search identifications using the Peaks *DeNovo*-assisted database search, we could expand identifications using MARS by a further 77,129 assigned sequences. We mapped 6,551 further peptides to human source proteins, 5,828 peptides with a single amino-acid substitution, and 3,980 peptides to non-canonical origins (Figure 4A). 60,770 peptides remained unmatched.

We then further stratified the obtained non-canonical HLAp according to the annotation of their origins as described in the last section (Figure 4B). Peptides uniquely mapped to protein coding regions formed the largest subcategory, comprising of 1150 HLAp, while 811 peptides remained unmatched. We found overall 1,423 peptides associated with lncRNA origin. In order to validate that the FDR of these identifications was as expected from the Chong *et al.* dataset, we again selected a set of peptides for spectral validation. We chose 19 peptide candidates that had a MARS score of >89.7 (region 1 and 2), and we further prioritized peptides if more than one sequence supported the identification of an lncRNA gene. We validated 11/19 spectra (42.2% FDR), which was in the range of the expected FDR of 39.5% from previous results for Region 1 and 2 (MARS score >89.7). For Region 1 only (MARS score >94.7) we again yielded high accuracy in annotation with 7/8 correct assignments and an FDR of 12.5% (Figure 4C, Table 1).

In summary, MARS identifies thousands of peptides in addition to peptides identified with standard database-search approaches at an accuracy of above 85.7%, as confirmed in this independent primary tumour dataset, paving the way for transcriptomics-independent antigen discovery workflows to inform the development of novel cancer immunotherapies.

A main limitation of non-canonical antigen discovery is the requirement of prior knowledge of expected protein sequences translated in a given sample. We here describe how we can rescue identification of non-canonical lncRNA-derived HLA-associated peptides from LC-MS/MS data without the requirement for integration of RNA sequencing data or, in fact, any other sample-specific sequence information other than HLA type, using a discovery engine that improves candidate selection from *de novo* sequencing results. Since in immunopeptidomics the lengths of peptides and positions of individual amino acids in the peptides are crucial for the HLA binding and T-cell response, we here used a study point-wise-full-sequence-recall (PSTR) to assess the development of MARS instead of improvements of single tag recall STR which assesses the percentage of individual amino acids being correctly identified [3, 34].

We evaluated the increased sensitivity using an external dataset from Wilhelm et al. [17], and confirmed an overall increase in performance of MARS *de novo* candidate selection in comparison to the best ALC scoring peptide as reported by Peaks *DeNovo* by over 7%. Interestingly, we saw outstanding performance for B27, B40 and B41 alleles. B27 has the requirement of an arginine at position two within the peptide sequence, while B40 and B41 are known to require a glutamic acid in this position. Hence, chemical properties of the peptide ligands are expected to be rather distinct, and we were not able to explain this much increased level of performance.

We did not observe a significant improvement for identification of HLAp originating from HLA-C. Since peptide data was acquired in equimolar ratios for the HLAp library datasets, this could not be due to lower expression of HLA-C in comparison to HLA-A and -B as generally observed across different datasets. Since NetMHCpan-4.1 performs equally well across all three loci, we conclude that most likely we did not evaluate a sufficient number of HLA-C peptide ligands to reach a significantly better performance.

We then proceeded to evaluate how many peptide identifications could be rescued from a list of published non-canonical HLAp that had been generated using a proteogenomics approach. MARS was

345 able to identify 45% of the reported sequences highlighting the power of MARS to identify non-canonical
346 antigens in the absence of transcriptomics and RiboSeq datasets.
347 One limitation remains the high false discovery rate associated with *de novo* sequencing results. With
348 MARS, the FDR could be further limited by assigning the peptide sequences to a possible origin in the
349 human genome and transcriptome, resulting in a manageable FDR range below 14.3% for MARS scores
350 above 94.8, compared to generally anticipated FDR of 65% for *de novo* sequencing results. Furthermore,
351 considering only those sequences that mapped to a genome origin resulted in a substantial decrease of
352 the FDR (to 0% in our experimental validation), demonstrating that MARS offers identification of non-
353 canonical peptide sequences at low FDR, and can therefore expand the peptide identification beyond
354 sophisticated proteogenomics approaches alone.
355 We here focused our analysis on non-canonical peptides derived specifically from lncRNA genes. The
356 question remains what the origin is of the parge proportion of non-assigned peptide sequences identified
357 by MARS. Post-translationally modified peptides, peptides derived from proteasomal fusion events, and
358 peptides of non-human origin are all potential candidates to explain this space, which remains to be
359 further investigated. Finally, the observed higher proportion of sequences with lower MARS score within
360 the unmatched peptides indicates a higher FDR in this subcategory, and further suggests that these
361 spectra may not contain sufficient information for accurate sequence matching. This observation suggests
362 that MARS performance could further be improved by measurement of higher quality spectra for each
363 peptide ion species, i.e. by increased ion accumulation times.
364 We could not validate any cysteine containing peptide assignments, confirming the known observed
365 underrepresentation of cysteines in immunopeptidomics data [35, 36], and highlighting the need for specific
366 modification of the otherwise highly reactive amino acid .
367 Finally, we were able to shortlist novel lncRNA-derived peptide candidates from primary human tumour
368 tissue. We shortlisted HLAp if more than one sequence supported the identification of a specific lncRNA

369 gene. We validated the sequence identifications with spectral matching to their synthetic counterpart,
370 and we validated the FDRs measured in our previous evaluation. Note that we did not set a threshold for
371 number of I and L in the peptide sequence, which could theoretically further reduce the FDR.
372 Finally, we identified and validated 11 peptides originating from lncRNA origins in human cervical cancer
373 tissue. Peptide KVHVFLVKKT, which originated from lncRNA HLA-F-AS1, plays an oncogenic role in both
374 colorectal cancer and breast cancer [37, 38]. The identified peptides MTMSTILSK and MTMSTILSKK were
375 derived from lncRNA lnckb.42318 also known as MIR4458HG (ENSG00000247516), which was reported
376 as highly expressed lncRNA in 407 TCGA evaluated ovarian tumors [39]. It has further been reported that,
377 together with CYTOR, and MAPTAS1, MIR4458HG was an independent prognostic factor in breast cancer
378 patients (https://doi.org/10.21203/rs.2.24062/v1, preprint), with high expression associated with lower
379 hazard ratios.
380 We conclude that MARS is a suitable platform for non-canonical antigen discovery in cancer
381 immunopeptidomics datasets in the absence of high quality RNA sequencing data. We further highlight
382 that the increases and unbiased results reported provide an exciting opportunity to expand non-canonical
383 cancer antigen discovery to post-translationally modified peptides, peptides derived from proteasomal
384 fusion events, and non-human sources, i.e. the microbiome.
385
386

METHODS

Human tissue material

Tissue samples were obtained from Tissue Solutions Ltd.. Ethical approval was granted by the Central University Research Ethics Committee (CUREC) of the University of Oxford under reference R68126/RE001.

Cell lines

All cell lines were mycoplasma-negative and cultured at 37 °C in the presence of 5% CO_2. The 721.221 HLA-class I deficient cell line that only expresses residual amounts of C*01:01 [40] transfected with CD4 (CD4.221) and selected HLA class I alleles (A*01:01, A*02:01, A*03:01, A*11:01, B*08:01, B*44:02, B*57:03, C*03:03 C*03:04) was kindly provided by Prof. Masafumi Takaguchi, Kumamoto University, Japan. Cells were grown in RPMI 1640 medium (Thermo Fisher) containing 10% fetal bovine serum (FBS), 2 mM l-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin (R10). R10 medium was supplemented with 0.15 mg/mL hygromycin B (Thermo Fisher) to maintain HLA-I expression as previously described [40, 41].

Baseline ALC Model, or One-Factor Model

Peaks *DeNovo* sequencing algorithm was performed using Peaks X (Bioinformatics Solutions) on each MS2 spectrum in the studied data to produce peptide-spectrum matches (PSMs). The quality of PSMs are measured by the average local confidence (ALC, represented by symbol *A* in the formula), which ranged in [0,100], with 100 being the best PSM quality. We selected PSMS with ALC in range [50, 100] to ensure reasonable input quality. The top-ALC peptide of each spectrum was used as its identity. No training was required for this one-factor model.

Integration of HLA Binding Affinity Ranking Factor (R): Two-Factor α Model

To obtain numerical characterization of HLA-binding affinities, The PSM candidates were input into NetMHCpan 4.1 model, and the output eluted ligand percentile rank (ERL, represented by symbol *R* in the formula) predictions were used to resemble the likelihood of HLA presentation of the corresponding peptides. The ELR predictions were considered robust across different biological samples and HLAs. The output ELR ranged between [0, 100], with the smaller ranks having higher likelihood of HLA presentation. We only selected PSMs with ERL in the range [0,10] to exclude non-binding peptides. By incorporating ELR into PSM candidates selection, a two-factor α model can be derived as: $(1-\alpha)A+\alpha(100-R)$ which is a positive affine combination of ALC (A) and ELR (R). The parameter α is in the range of [0,1],by the property of affine combination, the model outputs range is kept within [0,100] to ensure factors are normalized and of comparable scale. To obtain numerical characterization of HLA-binding affinities, The PSM candidates were input into NetMHCpan 4.1 model, and the output ELR (eluted ligand percentile rank) predictions used to resemble the likelihood of HLA presentation of the corresponding peptides.

Integration of the Retention Time Factor (T), Three-Factor MARS Model

We used an adapted version of DeepRTplus, which is a special type of deep neural network model termed capsule neural networks (CapsNets), to predict the RT of input peptides. Before the trained DeepRTplus model can be used for predicting into a new set of peptides, it needs to be recalibrated for each chromatographic setup, since peptides" RT observations are known to be chromatography dependent. We used the sequence-RT pairs from de-novo peptides with ALC > 95 as a small but reliable set of identified peptides to re-calibrate the pre-trained DeepRTplus model, forming a transfer learning strategy to increase prediction accuracy.

To develop the three-factor model by incorporating the difference in peptide chromatographic retention time (DRT, represented by symbol *T* in the formula) out of the DeepRTplus output as the input, firstly the

absolute difference between observed and predicted RT: $\Delta T = | T_{predicted} - T_{observed} |$ was computed. Secondly, T was computed as the ΔT s normalised to the interval of [0,100] to match the range of ALC and ELR: $T = 100(\Delta T_{max} - \Delta T)/(\Delta T_{max} - \Delta T_{min})$. By this normalisation, T was transformed to a reciprocal score range, with high values being better than low values. Thirdly, to take both ELR and DRT into selecting de-novo PSM candidates, a three-factor model was defined as: $(1-\alpha-\beta)A+\alpha(100-R)+\beta T$.

Performance Evaluation: *Fractional Ranks (F-Ranks)*

Fractional Ranks (F-Ranks) Fractional rank (F-Rank) of a correct peptide sequence for each spectrum is defined as the rank by descending order of this peptide divided by the total number of candidates of that spectrum. For peptides with tie scores, average rank was used as the numerator. It measures peptide identification performance improvement/deterioration at individual spectrum level. The reduction of mean F-rank of all correct peptides by multi-factor model outputs over the baseline ALC model was to suggest the promotion of the correct peptides. Differences in average F rank scores were evaluated with a paired T test.

Performance evaluation: Full-Sequence Recall (FSR)

FSR is defined as the ratio between the total number of correctly identified peptides and the total number of peptides in the given dataset. It measures the peptide identification sensitivity at the spectra ensemble level. For each factor model, the top-scoring PSM of each MS2 spectrum was used as an identified peptide sequence, whereas the database peptide sequence of the same MS2 spectrum was considered the correct peptide sequence. A peptide was considered correctly identified if the peptide predicted by the given model was identical (up to I/L indistinguishability) to the database peptide. Differences in FSR were evaluated with a paired T test.

Strategy for MARS candidate HLAp origin annotation and database curation

We coarsely stratified the HLAp origins into three categories: "human protein", "single amino acid substitution", and "non-canonical". Besides, we put all HLAp not strongly associated with these three categories into another "unmatched" category. Seqkit (Shen et al. 2016) was employed for locating MARS peptides in the curated databases. For these searches, we excluded all database search against SwissProt assigned peptides with a score cut-off of -10lgP>20, and searching the remaining *de novo* candidates as shortlisted by the MARS identification module.

Firstly, we searched Uniprot and Ensembl (only human protein entries) with isoforms as the "human proteome". MARS peptides matching exactly to human proteome were considered "Human Proteome" originated HLAp, matching human protein with 1AA difference were considered "1AA substitution". These two categories of peptides were considered of canonical origin, and excluded from non-canonical category. To further examine the possible alternative origins of the peptides of non-canonical origin, we used GENCODE v40 six-frame translated reference genome and three-frame translated reference transcriptome to form the protein database. All annotations were combined into Genecode "Biotypes" as follows: "Protein coding": protein_coding; "lncRNA": lncRNA; "other RNAs": miRNA, misc_RNA, Mt_rRNA, Mt_tRNA, ribozyme, rRNA, rRNA_pseudogene, scaRNA, scRNA, snoRNA, snRNA, sRNA, vault_RNA; "pseudogene": all biotypes including "pseudogene" (polymorphic, processed, transcribed_processed, transcribed_unitary, transcribed_unprocessed, translated_processed, translated_unprocessed, unitary, unprocessed); "IG": IG_C_gene, IG_C_pseudogene, IG_D_gene, IG_J_gene, IG_J_pseudogene, IG_pseudogene, IG_V_gene, IG_V_pseudogene; "TR": TR_C_gene, TR_D_gene, TR_J_gene , TR_J_pseudogene, TR_V_gene, TR_V_pseudogened and "TEC": TEC. We integrated the following databases for lncRNA mapping: LNCipedia [42] , lncRNAKB [43], and added a separate "TE" category using annotations from Dfam [44], and an Erv database [45]. We report our annotations in Upset plots to include multiple assignments.

483

484 **LC-MS acquisition**

485 All synthetic peptides were obtained from Genscript as a library service and crude purity. We measured
486 the peptide libraries by LC-MS on an Ultimate 3000 RSLCnano System coupled with an Q Exactive™ HF
487 Hybrid Quadrupole-Orbitrap™ Mass Spectrometer (Thermo Scientific). Peptides were loaded onto the
488 analytical column (PepMap C18 column, 2 µm particle size, 75 µm × 50 cm; Thermo Scientific) and eluted
489 in a 60 min linear gradient from 3% to 25% ACN in 1% DMSO/0.1% formic acid at a flow rate of 250 nl/min.
490 Peptides were introduced to the mass spectrometer using an EasySpray source at 2000 V and 45˚C, and
491 the transfer tube temperature was set to 305˚C. Mass spectrometry (MS) detection was performed with
492 a resolution of 120,000 for full MS (320-1600m/z scan range) and AGC target of 300,000. Precursors were
493 isolated using an isolation width of 1.6 amu and fragmented using high-energy collisional dissociation
494 (HCD) set at 25 for peptides with 2-4 charges. Fragments resolution was set at 60,000 with AGC target of
495 50000.

496

497 **LC-MS datasets and data analysis**

498 We utilized the following datasets: Data for the C1866 cell line (A*01:01, B*08:01, B*44:02, C*05:01, and
499 C*07:01) and for single allele transfectant cell lines is partially available in PXD015489 [40]. Data for the
500 Jurkat cell line (A*03:01, B*07:02, B*35:03, C*04:01, C07:02) is available in PXD011723 [46]. Wilhelm et al.
501 peptide standard datasets are available at PXD021013 [17], and Chong et al. datasets are available at
502 PXD013649 [2].
503 Peaks vX (Bioinformatics Solutions) was used to identify to search against a database containing the
504 synthetic standard sequences. For database search, we used no enzyme specificity (setting: unspecific).
505 The precursor mass error range was set within 5 ppm. Only oxidation was selected as variable
506 modification, unless otherwise indicated in the main text. For the *de novo* sequencing, the precursor mass

error range was set within 4.5 ppm, and only oxidation was selected as a variable modification unless otherwise indicated.

Spectral Matching

Spectral comparisons were performed manually, with the help of the Universal Spectrum Explorer (USE)[47]. If a synthetic peptide was not detected by LC-MS measurement, we considered the peptide sequence identification as "false", as it is equally unlikely that the identical peptide molecule was successfully ionized and detected during the experiment.

Trinity De-Novo RNA Assembly

Raw RNA sequencing reads had adaptors removed and low quality bases trimmed using Trim_Galore v0.6.2 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore) and quality control metrics assessed using FastQC v0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). A custom STAR (v2.7.3a) [48] genome reference was created utilising UCSC hg38 genome sequence. Reads were then mapped to this custom genome reference using STAR with two passes, sorted and indexed by Samtools v1.10 [49]. Duplicates were marked using GATK v4.1.7.0 [50]. Unmapped reads from STAR were used as inputs for Trinity v2.11.0 [51] for de novo transcript assembly.

526 **ACKNOWLEDGEMENT**

527 We thank Prof. Masafumi Takiguchi, Kumamoto University, Japan for provision of CD4-expressing cell lines
528 transfected with single HLA-I alleles. This work was funded by the NIH project #1R21AI153978-01 (LCE),
529 The Leona M. and Harry B. Helmsley Charitable Trust (RM), the European Association for the Study of
530 Diabetes FSD/JDRF/Lilly Programme on Type 1 Diabetes Research 2019 (NT), Cancer Research UK Cancer
531 Immunology project award C55884/A21045 (NT), and Cancer Research UK RadNet Centre Award
532 C6078/A28736 (NT). Some immunopeptidomic datasets employed for this study were generated as part
533 of studies supported by Medical Research Council (MRC) programme grant MR/K012037, grants from the
534 National Institutes of Health (UM1 AI 100645 and R01 AI 118549) and by AMED grant 17FK0410302H0003
535 for AIDS research (PB). The computational aspects of this research were supported by the Wellcome Trust
536 Core Award Grant Number 203141/Z/16/Z and the NIHR Oxford BRC. The views expressed are those of
537 the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. PB is a Jenner
538 Institute Investigator.

539

540 **DATA AND CODE AVAILABILITY**

541 Primary cervical tumour mass spectrometry data will be published in the Proteomics Identifications
542 Database (https://www.ebi.ac.uk/pride/) upon acceptance of the manuscript. Code will be provided to
543 academic researchers upon reasonable request.

544

545

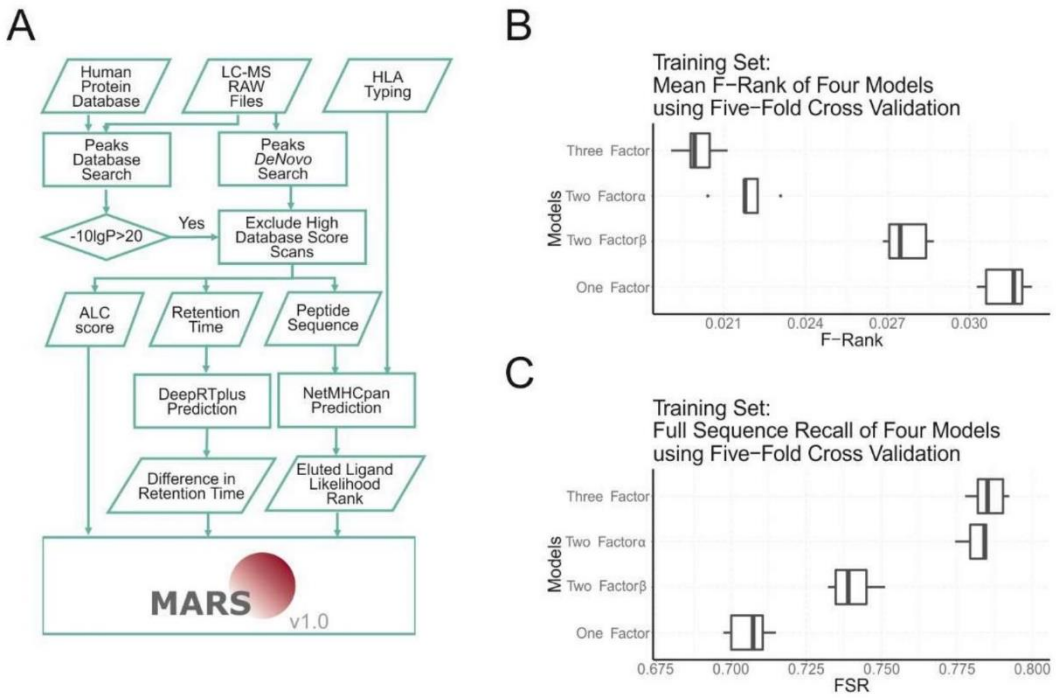
546 Table 1: lncRNA peptide validation

547 Red highlighted: MARS identification not reported by Peaks *DeNovo* as first candidate

PEPTIDE	MARS SCORE	HLA ALLELE	BINDING RANK	GENE LOCATION	LNC-RNA	ALT NAME	SAMPLE	SPECRAL MATCH
IIYFCLHKI	93.4	HLA-A*02:01	0.629	chr5:164989217-164989191	lnckb.44609	CTC-340A15.2	4,5	No
LPYIKWTI	93.6	HLA-B*51:01	0.076	chr5:164805831-164805808	lnckb.44609	CTC-340A15.2	7	Yes
LSSTLNKQI	90.1	HLA-C*02:02	1.143	chr5:164743628-164743602	lnckb.44609	CTC-340A15.2	1	No
IISASKVIL	92.3	HLA-A*02:01	0.617	chr11:90809963-90809989	lnckb.9710	DISC1FP1	7,9	No
TELKAWKI	95.2	HLA-B*3701	0.138	chr11:90693569-90693592	lnckb.9710	DISC1FP1	3	No
IRNVKIYLI	91.2	HLA-C*06:02	0.09		lnckb.45851	HLA-F-AS1	10	No
KVHVFLVKK	99.7	HLA-A*03:01	0.029	chr6:29750586-29750560	lnckb.45851	HLA-F-AS1	1,8	Yes
RQAPALRHL	92.5	HLA-C*06:02	0.133	chr13:94909331-94909357	lnckb.14510	LOC101927284	9	Yes
VIFSGIRSL	95.4	HLA-A*02:01	0.081	chr13:94926124-94926150	lnckb.14510	LOC101927284	5	Yes
VLLSIFIEHL	94.2	HLA-A*02:01	1.15	chr13:94845116-94845145	lnckb.14510	LOC101927284	7	No
ALRAVTLTAK	97.9	HLA-A*03:01	0.099	chr3:35175054-35175080	lnckb.36918	LOC101928135	8,10	Yes
KQLNKQLIK	98.3	HLA-A*0301	0.053	chr8:78259721-78259747	lnckb.53745	LOC105375911	4,9	Yes
VTIFQNRVK	92	HLA-A*11:01	0.575	chr8:78413273-78413247	lnckb.53745	LOC105375911	8	Yes
KTKIITGSK	94.1	HLA-A*03:01	0.192	chrX:75675106-75675132	lnckb.58303	LOC107985664	8	Yes
SLSILSLKV	92.9	HLA-B*13:02	0.164	chrX:75543586-75543560	lnckb.58303	LOC107985664	1	No
IPHGEIPDTSA	94.9	HLA-B*56:01	0.064	chr2:172323392-172323424	lnckb.31346	LOC107985960	5	Yes
LPNFARII	92.6	HLA-B*07:02	0.29	chr2:172280278-172280255	lnckb.31346	LOC107985960	4,8	No
MTMSTILSK	99.2	HLA-A*11:01	0.008	chr5:8460009-8460035	lnckb.42318	MIR4458HG	8	Yes
MTMSTILSKK	99.8	HLA-A*03:01	0.199	chr5:8460009-8460038	lnckb.42318	MIR4458HG	4,8,9	Yes

548

Figure 1



549

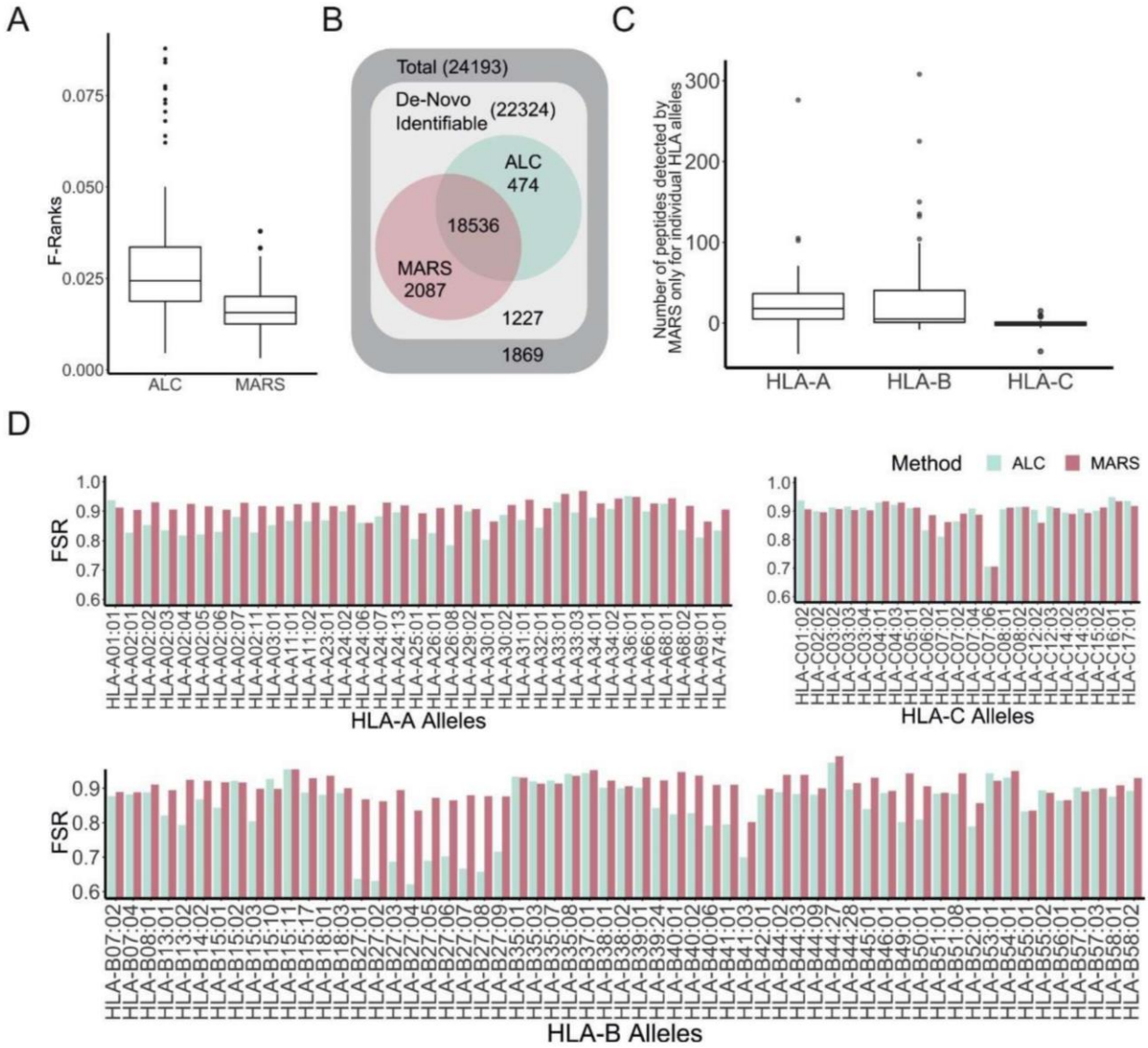
550 **Figure 1: Schematic of MARS workflow and performance evaluation in a five-fold cross-validation.** (A) Schematic

551 of integrative modules in the MARS pipeline (B) Mean F-rank score of all partitions, and (C) Full sequence recall (FSR)

552 achieved across the partitions in the target peptide space.

1

Figure 2



553

554 **Figure 2: MARS sensitivity evaluation in *Wilhelm et al.* large HLA peptide library datasets.** (A) The average F-rank

555 distribution across all five partitions and (B) Venn diagram depicting the number of peptide sequences identified by

556 ALC (green) and MARS (red), respectively, within those sequences that were present in the 20 *de novo* candidates

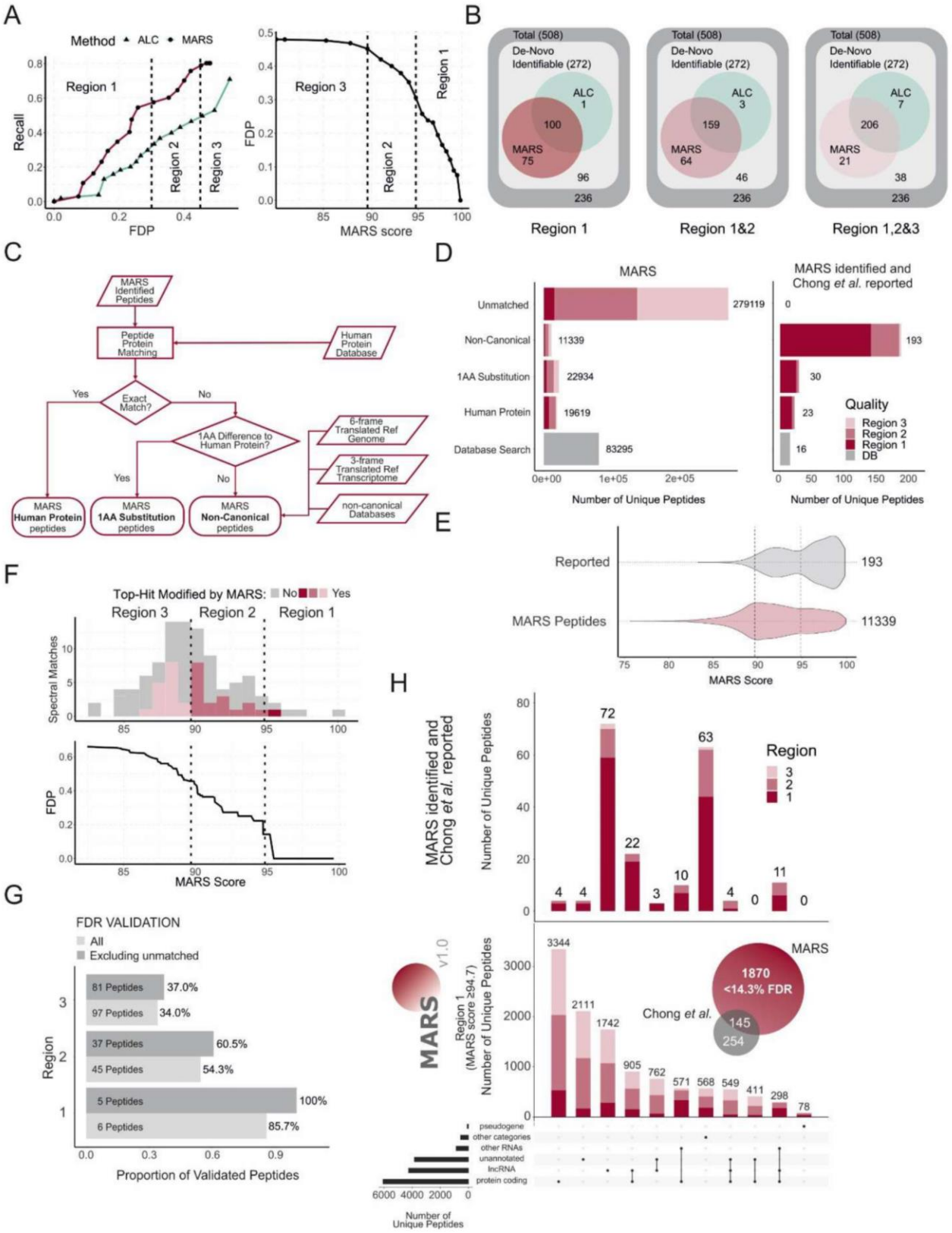
557 as reported by Peaks ("*de novo* identifiable", dark grey). Total sequence space is shown in light grey. (C) Box plots

558 indicating the number of peptide sequences uniquely identified by MARS for all HLA-A, -B, and -C alleles combined

559 (D) Full sequence recall (FSR) for individual HLA alleles for ALC (green) and MARS (red).

560

Figure 3


A
Method ▲ ALC ● MARS
Region 1
Region 2
Region 3
Recall
FDP
MARS score
B
Total (508)
De-Novo Identifiable (272)
ALC 1
100
MARS 75
96
236
Region 1
ALC 3
159
MARS 64
46
Region 1&2
ALC 7
206
MARS 21
38
Region 1,2&3
C
MARS Identified Peptides
Peptide Protein Matching
Human Protein Database
Exact Match?
Yes
No
1AA Difference to Human Protein?
6-frame Translated Ref Genome
3-frame Translated Ref Transcriptome
non-canonical Databases
MARS Human Protein peptides
MARS 1AA Substitution peptides
MARS Non-Canonical peptides
D
MARS
Unmatched 279119
Non-Canonical 11339
1AA Substitution 22934
Human Protein 19619
Database Search 83295
Number of Unique Peptides
MARS identified and Chong et al. reported
0
193
30
23
16
Quality
Region 3
Region 2
Region 1
DB
E
Reported 193
MARS Peptides 11339
MARS Score
F
Top-Hit Modified by MARS: No Yes
Spectral Matches
FDP
MARS Score
G
FDR VALIDATION
All
Excluding unmatched
81 Peptides 37.0%
97 Peptides 34.0%
37 Peptides 60.5%
45 Peptides 54.3%
5 Peptides 100%
6 Peptides 85.7%
Region
Proportion of Validated Peptides
H
MARS identified and Chong et al. reported
Number of Unique Peptides
Region
MARS v1.0
Region 1 (MARS score ≥94.7)
1870 <14.3% FDR
Chong et al.
145
254
pseudogene
other categories
other RNAs
unannotated
lncRNA
protein coding
Number of Unique Peptides

561

562 **Figure 3: Non-canonical peptide identification using MARS in a published dataset from Chong *et al.* [2]**. (A) Left panel:
563 Sequence recall for lncRNA peptide sequences identified by Chong at all across all samples plotted against the false
564 discovery rate. The three regions are defined by FDR cut-offs at 30% and 45%, respectively. The right panel shows
565 these regions in relation to the MARS score. (B) Venn diagrams demonstrating the overlap of sequence
566 identifications for each of the three defined regions from (A). (C) Overview of MARS search engine and blast search
567 strategy for identification of source origin. (D) Left panel: Bar graph showing the categorization of the MARS hits as
568 indicated. Right panel: Mutually identified peptides by MARS and Chong *et al.* from the 508 reported non-canonical
569 identifications assigned to the MARS categories as indicated. Colored subsections of the graph indicate the original
570 alternative classification as applied by *Chong et al.* in the original publication (E) Score distribution of all MARS
571 identified non-canonical peptides (11339) in comparison to those mutually identified by Chong *et al.* (193). (F)
572 Experimental validation of the FDR using a synthetic peptide standard. Histogram summarizing the peptide score
573 distribution of the synthetic peptide standard. Shown are number of peptides identified by MARS only for indicated
574 MARS score bins (upper panel), and FDP progression in relation to declining MARS score (lower panel) calculated on
575 the spectral matching outcome. (G) Number and proportion of all confirmed peptides for all synthetic standard
576 sequences and for the subset of peptides that were mapped to the human genome ("excluding unmatched"). (H)
577 Upset plot of the MARS non-canonical peptide top 10 categories for all hits in region 1, 2, and 3 (lower panel), and
578 mutual identifications as reported in Chong *et al.* (upper graph). The Venn diagram indicates the overall number of
579 non-canonical sequences that were identified with MARS at a FDR ≤ 14.3%.
580

Figure 4

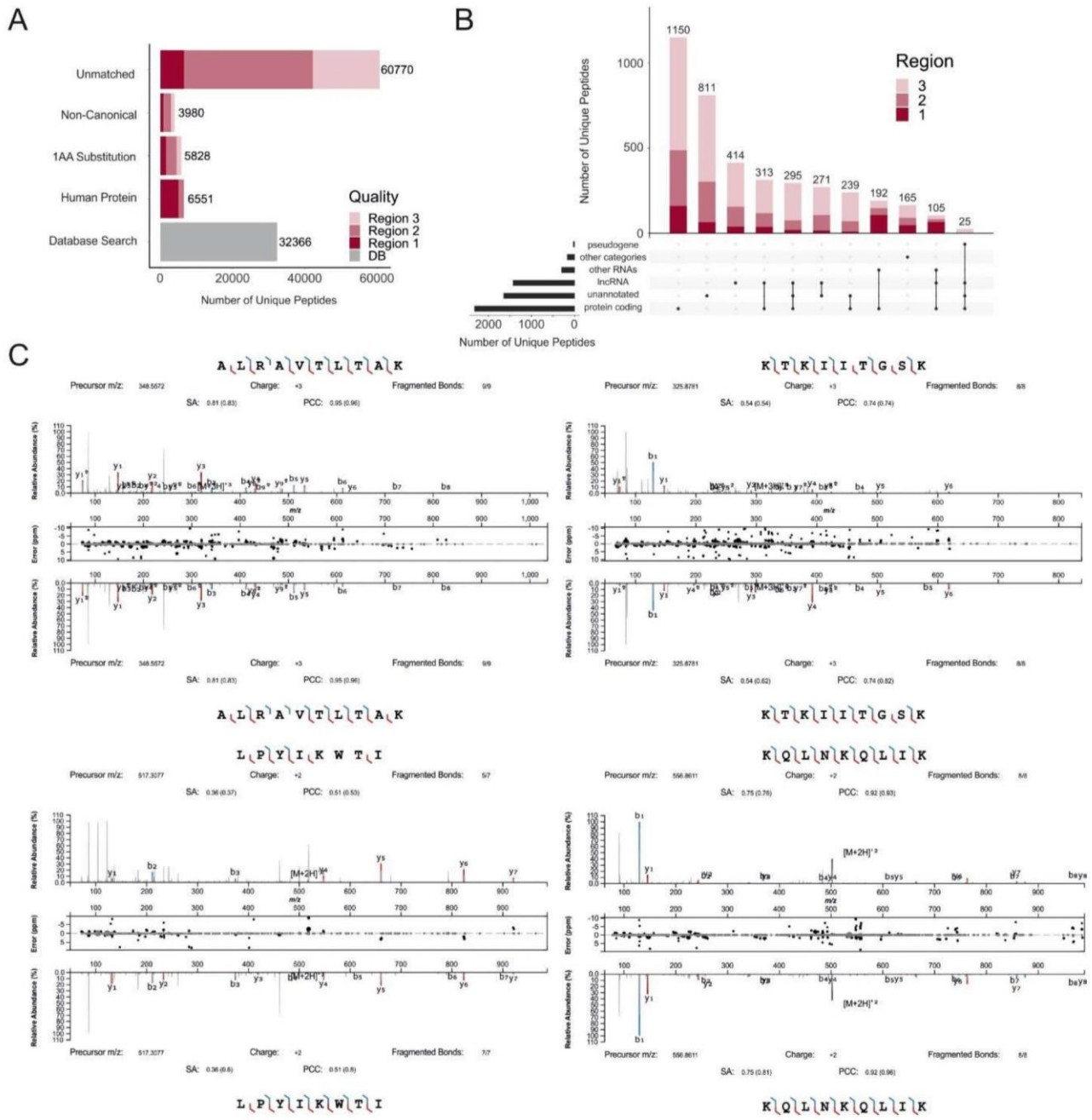


Figure 4: Identification of novel non-canonical HLAp in primary tumour tissue using MARS. (A) Bar graph showing the categorization of the MARS hits as indicated. (B) Upset plot of the MARS non-canonical peptide category for all hits (C) Bar graph depicting the level of evidence found in trinity assemblies of RNAseq data for each lncRNA category. (D) Score distribution of all MARS peptide hits in comparison of those that are validated with additional RNA sequence evidence. (E) Spectral matching validation for indicated sequences identified in the primary tumour tissue.

REFERENCES

1. Jaeger, A.M. et al. Deciphering the immunopeptidome in vivo reveals new tumour antigens. *Nature* (2022).
2. Chong, C. et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun* **11**, 1293 (2020).
3. Muth, T. & Renard, B.Y. Evaluating de novo sequencing in proteomics: already an accurate alternative to database-driven peptide identification? *Brief Bioinform* **19**, 954-970 (2018).
4. Bassani-Sternberg, M. et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat Commun* **7**, 13404 (2016).
5. Ternette, N. et al. Defining the HLA class I-associated viral antigen repertoire from HIV-1-infected human cells. *Eur J Immunol* **46**, 60-69 (2016).
6. Bettencourt, P. et al. Identification of antigens presented by MHC for vaccines against tuberculosis. *NPJ Vaccines* **5**, 2 (2020).
7. Attig, J. et al. LTR retroelement expansion of the human cancer transcriptome and immunopeptidome revealed by de novo transcript assembly. *Genome Res* **29**, 1578-1590 (2019).
8. Kracht, M.J. et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med* **23**, 501-507 (2017).
9. Zanker, D.J. et al. Influenza A Virus Infection Induces Viral and Cellular Defective Ribosomal Products Encoded by Alternative Reading Frames. *J Immunol* **202**, 3370-3380 (2019).
10. Parker, R. et al. The Choice of Search Engine Affects Sequencing Depth and HLA Class I Allele-Specific Peptide Repertoires. *Mol Cell Proteomics* **20**, 100124 (2021).
11. Bichmann, L. et al. MHCquant: Automated and Reproducible Data Analysis for Immunopeptidomics. *J Proteome Res* **18**, 3876-3884 (2019).
12. Mann, M. & Wilm, M. Error-tolerant identification of peptides in sequence databases by peptide sequence tags. *Anal Chem* **66**, 4390-4399 (1994).
13. Yates, J.R., 3rd, Eng, J.K., McCormack, A.L. & Schieltz, D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* **67**, 1426-1436 (1995).
14. Nesvizhskii, A.I. Protein identification by tandem mass spectrometry and sequence database searching. *Methods Mol Biol* **367**, 87-119 (2007).
15. AS, C.S., Bouwmeester, R., Martens, L. & Degroeve, S. Accurate peptide fragmentation predictions allow data driven approaches to replace and improve upon proteomics search engine scoring functions. *Bioinformatics* **35**, 5243-5248 (2019).
16. Gessulat, S. et al. Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods* **16**, 509-518 (2019).
17. Wilhelm, M. et al. Deep learning boosts sensitivity of mass spectrometry-based immunopeptidomics. *Nat Commun* **12**, 3346 (2021).
18. Kong, A.T., Leprevost, F.V., Avtonomov, D.M., Mellacheruvu, D. & Nesvizhskii, A.I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat Methods* **14**, 513-520 (2017).

19. Li, K., Jain, A., Malovannaya, A., Wen, B. & Zhang, B. DeepRescore: Leveraging Deep Learning to Improve Peptide Identification in Immunopeptidomics. *Proteomics* **20**, e1900334 (2020).
20. Tran, N.H., Zhang, X., Xin, L., Shan, B. & Li, M. De novo peptide sequencing by deep learning. *Proc Natl Acad Sci U S A* **114**, 8247-8252 (2017).
21. Yang, H., Chi, H., Zeng, W.F., Zhou, W.J. & He, S.M. pNovo 3: precise de novo peptide sequencing using a learning-to-rank framework. *Bioinformatics* **35**, i183-i190 (2019).
22. Ma, B. et al. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid Commun Mass Spectrom* **17**, 2337-2342 (2003).
23. Erhard, F., Dolken, L., Schilling, B. & Schlosser, A. Identification of the Cryptic HLA-I Immunopeptidome. *Cancer Immunol Res* **8**, 1018-1026 (2020).
24. Laumont, C.M. et al. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun* **7**, 10238 (2016).
25. Ouspenskaia, T. et al. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat Biotechnol* **40**, 209-217 (2022).
26. Andreatta, M. et al. MS-rescue: A Computational Pipeline to Increase the Quality and Yield of Immunopeptidomics Experiments. *Proteomics*, e1800357 (2018).
27. Ma, C. et al. Improved Peptide Retention Time Prediction in Liquid Chromatography through Deep Learning. *Anal Chem* **90**, 10881-10888 (2018).
28. Tran, N.H. et al. Personalized deep learning of individual immunopeptidomes to identify neoantigens for cancer vaccines. *Nat Mach Intell* **2** (2020).
29. Zhang, J. et al. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **11**, M111 010587 (2012).
30. Tran, N.H. et al. Deep learning enables de novo peptide sequencing from data-independent-acquisition mass spectrometry. *Nat Methods* **16**, 63-66 (2019).
31. Gonzalez-Galarza, F.F. et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. *Nucleic Acids Res* **48**, D783-D788 (2020).
32. Martini, S., Nielsen, M., Peters, B. & Sette, A. The Immune Epitope Database and Analysis Resource Program 2003-2018: reflections and outlook. *Immunogenetics* **72**, 57-76 (2020).
33. Vita, R. et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res* **47**, D339-D343 (2019).
34. Xu, C. & Ma, B. Complexity and scoring function of MS/MS peptide de novo sequencing. *Comput Syst Bioinformatics Conf*, 361-369 (2006).
35. Meadows, L. et al. The HLA-A*0201-restricted H-Y antigen contains a posttranslationally modified cysteine that significantly affects T cell recognition. *Immunity* **6**, 273-281 (1997).
36. Trujillo, J.A. et al. The cellular redox environment alters antigen presentation. *J Biol Chem* **289**, 27979-27991 (2014).
37. Huang, Y. et al. HLA-F-AS1/miR-330-3p/PFN1 axis promotes colorectal cancer progression. *Life Sci* **254**, 117180 (2020).
38. Wu, D., Jia, H., Zhang, Z. & Li, S. STAT3-induced HLA-F-AS1 promotes cell proliferation and stemness characteristics in triple negative breast cancer cells by upregulating TRABD. *Bioorg Chem* **109**, 104722 (2021).
39. Akrami, R. et al. Comprehensive analysis of long non-coding RNAs in ovarian cancer reveals global patterns and targeted DNA amplification. *PLoS One* **8**, e80306 (2013).
40. Partridge, T. et al. Discrimination Between Human Leukocyte Antigen Class I-Bound and Co-Purified HIV-Derived Peptides in Immunopeptidomics Workflows. *Front Immunol* **9**, 912 (2018).
41. Paes, W. et al. Contribution of proteasome-catalyzed peptide cis-splicing to viral targeting by CD8(+) T cells in HIV-1 infection. *Proc Natl Acad Sci U S A* **116**, 24748-24759 (2019).

683 42. Volders, P.J. et al. LNCipedia 5: towards a reference set of human long non-coding RNAs.
684 *Nucleic Acids Res* **47**, D135-D139 (2019).
685 43. Seifuddin, F. et al. lncRNAKB, a knowledgebase of tissue-specific functional annotation
686 and trait association of long noncoding RNA. *Sci Data* **7**, 326 (2020).
687 44. Storer, J., Hubley, R., Rosen, J., Wheeler, T.J. & Smit, A.F. The Dfam community resource
688 of transposable element families, sequence models, and genome annotations. *Mob DNA*
689 **12**, 2 (2021).
690 45. Saini, S.K. et al. Human endogenous retroviruses form a reservoir of T cell targets in
691 hematological cancers. *Nat Commun* **11**, 5660 (2020).
692 46. Nicastri, A., Liao, H., Muller, J., Purcell, A.W. & Ternette, N. The Choice of HLA-Associated
693 Peptide Enrichment and Purification Strategy Affects Peptide Yields and Creates a Bias
694 in Detected Sequence Repertoire. *Proteomics* **20**, e1900401 (2020).
695 47. Schmidt, T. et al. Universal Spectrum Explorer: A Standalone (Web-)Application for Cross-
696 Resource Spectrum Comparison. *J Proteome Res* **20**, 3388-3394 (2021).
697 48. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21
698 (2013).
699 49. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-
700 2079 (2009).
701 50. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing
702 next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).
703 51. Grabherr, M.G. et al. Full-length transcriptome assembly from RNA-Seq data without a
704 reference genome. *Nature Biotechnology* **29**, 644-U130 (2011).
705

706

SUPPLEMENTARY INFORMATION

Supplementary Figure 1: I replacement by L maintains validity of NetMHCpan 4.1 prediction across distinct HLA alleles

Supplementary Figure 2: HLA allele distributions for Chong et al. dataset, and cervical tumour patient cohort.

Supplementary Figure 3 (provided separately): **Spectral matching for FDR determination**

Supplementary Data Table 1: Five-Fold Cross-Validations using Baseline ALC model, Two-Factor alpha model, Two-Factor beta model, and Three-Factor MARS model. Column 2-5 shows the optimal parameters learned by grid search for each part, suggesting the models are robust. Column 6-8 gives the results of Paired Wilcoxon signed-rank tests on F-ranks of true peptide sequences; significant drop of F-ranks indicates improved identification power.

Supplementary Data Table 2: Number of peptide sequences reported to bind to each respective HLA allele in the Wilhelm et al. dataset

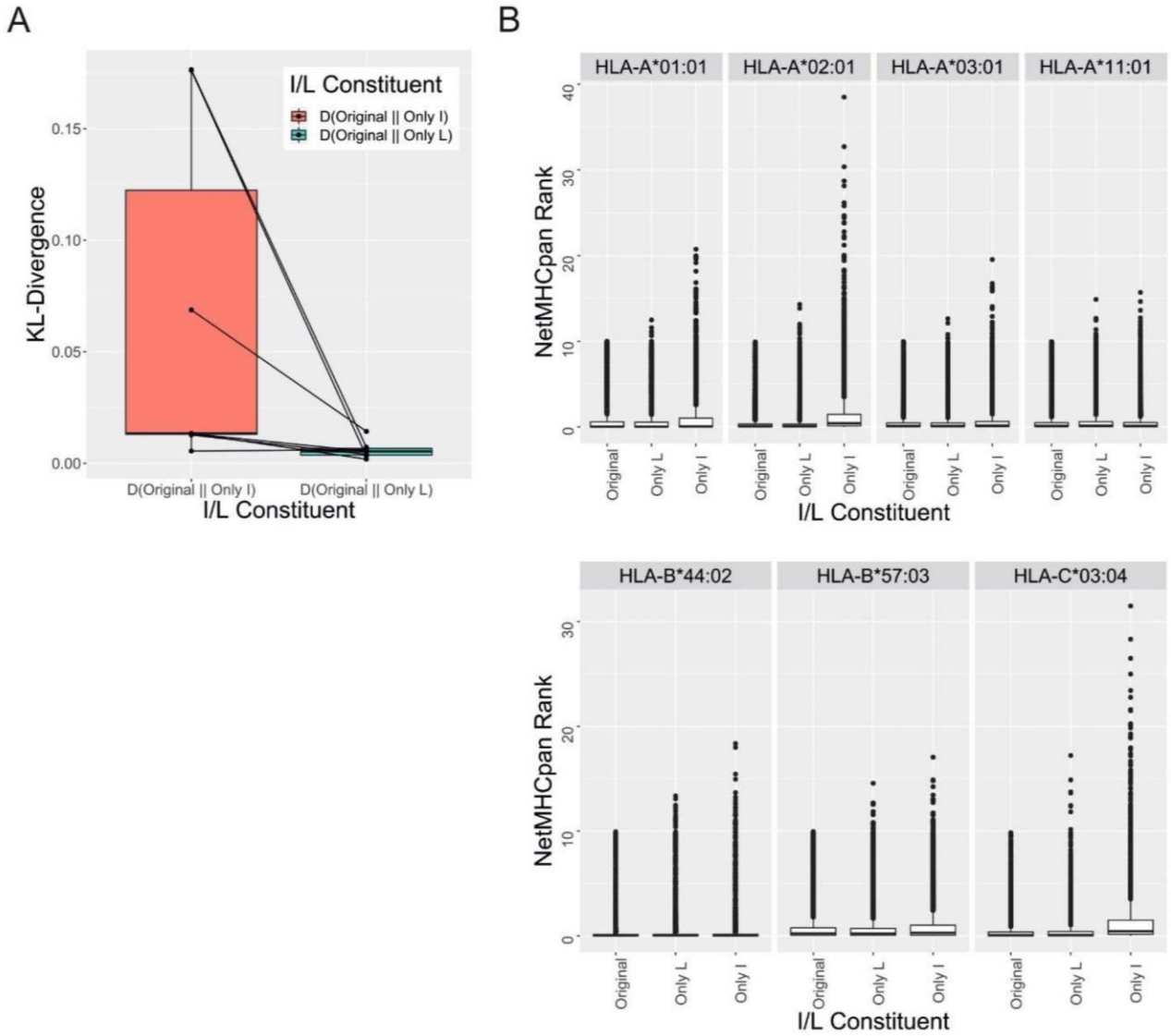
Supplementary Data Table 3: FSR rates for Chong et al. data for each sample

Supplementary Data Table 4: Overview of MARS peptide identifications for each sample in the Chong et al. dataset

Supplementary Data Table 5: HLA-typing results for cervical tumour patients

Supplementary Data Table 6: FDR determination: Peptide Validation Summary

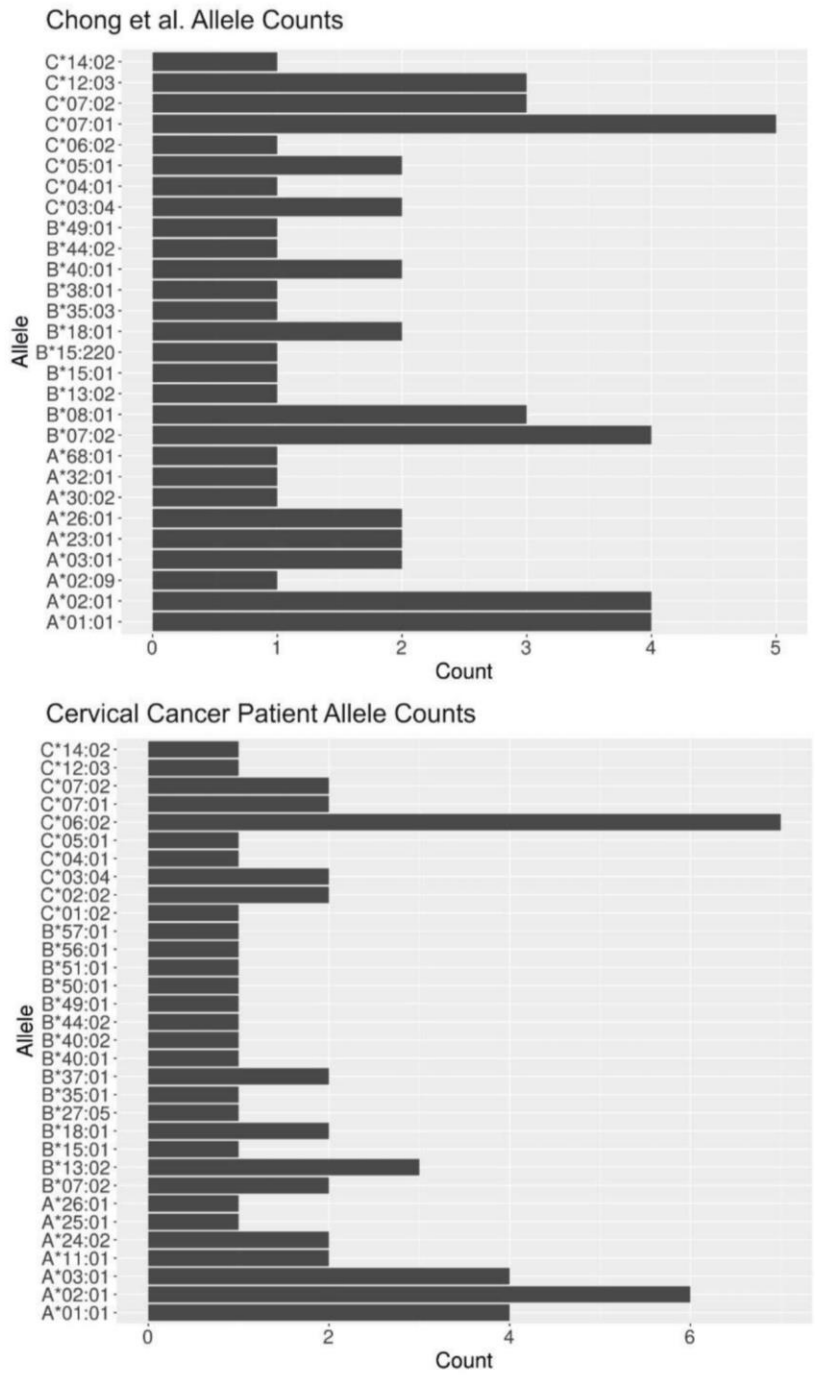
Supplementary Figure 1



733

734 **Supplementary Figure 1: I replacement by L maintains validity of NetMHCpan 4.1 prediction across**
735 **distinct HLA alleles.** (A) K-L divergences were computed as statistics to characterize the similarity of two
736 distributions D(Original || Only I) and D(Original || Only L). KLDiv_Boxplot shows that Only L had
737 consistently very small K-L divergences, and a paired Wilcoxon signed rank test suggested that D(Original
738 || Only I) is significantly greater than D(Original || Only L) with p = 0.0156. (B) Rank score distributions
739 for example alleles for original peptide sequence, all I replaced by L (Only L), and vice versa (Only I).

Supplementary Figure 2



Supplementary Figure 2: HLA allele distributions for Chong et al. dataset, and cervical tumour patient cohort.

Annex 6: main Python pipeline developed for HLA-peptidomics analysis

All related scripts can be found on: https://gitfront.io/r/acarre/bsHLg1kHCSRn/HLA-peptidomics/

The main function to analyze HLA-I peptidomics data of ECN90 β cells exposed or not to IFN-α is below.

```
import os
import pandas as pd
import src.toolbox.initialization as init
import src.toolbox.reformat_entering as reformat
import src.toolbox.length_filter as length
import src.toolbox.pepOcc_geneAnnot as occAnnot
import src.toolbox.ubiquity_filter as ubi
import src.toolbox.enrichment_filter as enrich
import src.toolbox.concatenation as concat
import src.toolbox.netMHC_I_pipeline as MHC
import src.toolbox.AAoccurence as AA
import numpy as np

os.system('echo $PATH')
#ROOT_PATH = "D:/"
ROOT_PATH = "/mnt/109E1E2A9E1E08BC/"
#ROOT_PATH = "/media/alexia/"
#FILE_PATH = "/media/alexia/DataAlexia/Projets/Scripts/HLA-I/IFNa_dataset/"
FILE_PATH = ROOT_PATH + "DataAlexia/Projets/Scripts/HLA-I/IFNa_dataset/"
INPUT_FILE_PATH = FILE_PATH + "Input/"
OUTPUT_FILE_PATH = FILE_PATH + "OutputNew_check4.1a/"

"""STEP 0: manually remove and isolate peptides explained only by a mRNA variant"""

"""INITIALIZATION : creation of the needed dictionaries"""
print("Generating the needed dictionaries...")
# Reading of files for dictionaries generation
annotationFile = pd.read_csv(INPUT_FILE_PATH + 'mainAnnot_homo_sapiens.txt',
                sep='\t',
                dtype = {
                  "Gene name":"str",
                  "UniProt accession" : "str"
                }
                )
atlasFile = pd.read_csv(INPUT_FILE_PATH + 'proteinatlas.tsv',
            sep='\t',
            dtype = {
              "Gene" : 'str',
              "Gene description" : "str",
              "Gene synonym" : "str",
              "RNA TS TPM" : "str",
              "RNA tissue category" : "str"
            }
            )
```

```
HPRDFile = pd.read_csv(INPUT_FILE_PATH + 'TISSUE_EXPRESSIONS.txt',
            sep='\t',
            dtype = {
              "GeneSymbol" : 'str',
              "Expression_term" : "str"
            }
            )

normalTissueFile = pd.read_csv(INPUT_FILE_PATH + "normal_tissue.tsv",
                sep = '\t',
                dtype={
                  "Gene name" : 'str',
                  "Tissue" : 'str',
                  'Level' : 'str'
                }
                )

def processDict(annotationFile, atlasFile, HPRDFile, normalTissueFile):
#Import files as dfs
  annotations = pd.DataFrame(annotationFile)
  atlas = pd.DataFrame(atlasFile)
  HPRDdf = pd.DataFrame(HPRDFile)
  normalTissuedf = pd.DataFrame(normalTissueFile)

#Apply processing functions = step 1 of the pipeline
  annotationDict, RNATSTPMDict, geneSynonymDict, geneDescriptionDict, expTermDict,
normalTissueDict = init.createDictionaries(annotations, atlas, HPRDdf, normalTissuedf)
  expTermDict = init.reformatDictValue(expTermDict)

  return annotationDict, RNATSTPMDict, geneSynonymDict, geneDescriptionDict, expTermDict,
normalTissueDict


"""REFORMATTING ENTERING FILE"""
def reformat_entering_data(data, type):
  data = reformat.reformatAccession(data, type)
  data = data[data.Accession != 0]
  data = reformat.reduceTable(data)

  return data

"""FILTERING FOR BETA CELL ENRICHMENT"""
def betaCell_enrichment(data, ubidf, ATHMdf, sandbergdf, geneDescriptionDict, expTermDict,
normalTissueDict, type):
  data = enrich.expTermExpression(data, expTermDict)
  data = enrich.addNormalTissue(data, normalTissueDict)
  data = enrich.reformatRNATSTPM_expTerm(data) #reformats only RNATSTPM col
  data = enrich.litEvidence(data)
  data = enrich.RNAevidence(data)
  data = enrich.protEvidence(data)
  data, ubidf = enrich.RNAprotLitEvidenceFilter(data, ubidf, type)
  data, ubidf = enrich.filterSandberg(data, ubidf, sandbergdf, type)
  data = enrich.addGeneDescri(data, geneDescriptionDict)
  data, ATHMdf = enrich.removeATHM(data, ATHMdf)

  return data, ubidf, ATHMdf

"""CONCATENATION OF RESULTS TABLES"""
def concatAndClean(allFormattedData):
  dfBasal = allFormattedData[0]
  dfIFNa = allFormattedData[1]

  finalTable = concat.concatResults(dfBasal, dfIFNa)
```

```
  finalTable = concat.cleanResults(finalTable)

  return finalTable

"""GENERTING MHC PREDICTIONS"""

allelesList = ["HLA-A02:01", "HLA-A03:01", "HLA-B40:01", "HLA-B49:01", "HLA-C03:04", "HLA-
C07:01", "HLA-E01:01"]

def MHCPredictionGeneration(data, filePath, type, allelesList):
  MHC.computeBindingPredictions(data, filePath, CURRENT_TYPE, allelesList)
  restrictiondfList = [] # initializes empty list

  first3colToName = ["Pos", "Peptide", "ID"]
  first2colToDrop = ["Pos", "ID"]
  last2col = ["Ave", "NB"]
  HLAcolNameList = ["Rank_" + s for s in allelesList]
  repeatedCol = ["core", "icore", "EL-score"]

  preColNames = []
  preColtoDrop = []

  for i in range(0, len(HLAcolNameList)):
    preColNames.extend(repeatedCol)
    preColNames.append(HLAcolNameList[i])

    preColtoDrop.extend(repeatedCol)

  first3colToName.extend(preColNames)
  colNames = first3colToName
  colNames.extend(last2col)

  first2colToDrop.extend(preColtoDrop)
  colsToDrop = first2colToDrop
  colsToDrop.append(last2col[0])

  for j in range(7):
    predictionFiles = pd.read_csv(OUTPUT_FILE_PATH + '{}mersPredictions{}.txt'.format(8 + j,
type), sep='\t', header=[1])

    predictiondf = pd.DataFrame(predictionFiles)
    predictiondf.columns = colNames
    predictiondf = predictiondf.drop(columns=colsToDrop)
    for col, content in predictiondf.iteritems():
      if "Rank_" in str(col):
        predictiondf = predictiondf.astype(
          { col: np.float64 }
        )
    restrictiondf = MHC.bindingPredictionsTreatment(predictiondf, allelesList)
    restrictiondfList.append(restrictiondf)

  concatRestrictiondf = pd.concat(restrictiondfList)

  dataAssigned = MHC.finalTableHLA_Assignment(data, concatRestrictiondf, OUTPUT_FILE_PATH)
  return dataAssigned


#if __name__ == "main":
"""PIPELINE : FILTERING STEPS"""
#Call of the function to generate the dictionaries
annotationDict, RNATSTPMDict, geneSynonymDict, geneDescriptionDict, expTermDict,
normalTissueDict = processDict(annotationFile, atlasFile, HPRDFile, normalTissueFile)
```

```
"""Consideration of datatypes """
#Takes into account both datatypes
datatype = ["Basal", "IFNa"]
allFormattedData = [] #initializes empty list to keep all results table
for type in datatype:
  global CURRENT_TYPE # defines a global variable for type
  CURRENT_TYPE = type
  dataFile = pd.read_csv(INPUT_FILE_PATH + 'DB search psm_{}.csv'.format(CURRENT_TYPE),
              sep=',',
              dtype = {
                "Peptide" : "str",
                "Length" : np.int64,
                "Fraction" : np.int64,
                "Accession" : "str",
                "PTM" : "str"
              }
              )
  sandbergFile = pd.read_csv(INPUT_FILE_PATH + 'Sandberg_cellType_enrichment.txt',
                sep='\t',
                dtype = {
                  "Genes": 'str',
                  "sampleID": 'str',
                  "gamma cell": np.float64,
                  "alpha cell": np.float64,
                  "beta cell": np.float64,
                  "acinar cell": np.float64,
                  "epsilon cell": np.float64,
                  "ductal cell": np.float64,
                  "delta cell": np.float64,
                  "Cell enrichment" : "str"
                }
                )

  print("\nRUNNING THE ANALYSIS OF THE {} FILE".format(CURRENT_TYPE))

  # Import files as dfs
  data = pd.DataFrame(dataFile)
  sandbergdf = pd.DataFrame(sandbergFile)

  data = data.fillna(0) # pandas cannot deal with empty cells, fill all with 0

  #FUNCTIONS CALLING
    # call of function to reformat data to enter pipeline
  data = reformat_entering_data(data, CURRENT_TYPE)
  data.to_csv(path_or_buf=OUTPUT_FILE_PATH + 'DB search
psm_{}_Entering.txt'.format(CURRENT_TYPE),
              index=False, sep='\t')
  # counts the AA occurences of all peptides and saves it into a specific file
  countAll = AA.countAAoccurence(data)
  countPercentageAll = AA.computeAAocc(countAll)
  countPercentageAll.to_csv(path_or_buf=OUTPUT_FILE_PATH +
'{}_AAcounts_all.txt'.format(CURRENT_TYPE),
                index=False, sep='\t')
  # call of function to filter on length of peptide
  data = length.lengthFilter(data, 8, 14, CURRENT_TYPE)
  # counts the AA occurences of all peptides and saves it into a specific file
  countLengthFiltered = AA.countAAoccurence(data)
  countPercLengthFiltered = AA.computeAAocc(countLengthFiltered)
  countPercLengthFiltered.to_csv(path_or_buf=OUTPUT_FILE_PATH +
'{}_AAcounts_LengthFiltered.txt'.format(CURRENT_TYPE),
                index=False, sep='\t')
  #calls the function to add "raw peptide" column
  data = concat.rawPeptide(data)
```

```
  #call of function to generate MHC binding predictions
  data = MHCPredictionGeneration(data, OUTPUT_FILE_PATH, CURRENT_TYPE, allelesList)
  #call of function to add peptide occurence AND gene annotation
  data = occAnnot.countOccurences(data, CURRENT_TYPE)
  data = occAnnot.annotGene(data, annotationDict)
  data.to_csv(path_or_buf=OUTPUT_FILE_PATH + '{}_Length_Filtered.txt'.format(CURRENT_TYPE),
            index=False, sep='\t')
  #call of function to filter ubi from non ubi peptides

  data = ubi.addGeneSyno(data, geneSynonymDict)
  data = ubi.nonUbiExpression(data, RNATSTPMDict)

  #call of function to filter peptides enriched in beta cells
  ubidf = pd.DataFrame()
  ATHMdf = pd.DataFrame()
  formattedData, ubidf, ATHMdf = betaCell_enrichment(data, ubidf, ATHMdf, sandbergdf,
geneDescriptionDict, expTermDict, normalTissueDict, type)

  allFormattedData.append(formattedData)

  #saves intermediate datframes as csv
  formattedData.to_csv(path_or_buf= OUTPUT_FILE_PATH + 'DB search
psm_{}_Formatted.txt'.format(CURRENT_TYPE), index=False, sep='\t')
  ubidf.to_csv(path_or_buf=OUTPUT_FILE_PATH +
'{}_ubiquitous_peptides.txt'.format(CURRENT_TYPE), index=False, sep='\t')
  ATHMdf.to_csv(path_or_buf=OUTPUT_FILE_PATH + '{}_ATHM_peptides.txt'.format(CURRENT_TYPE),
index=False, sep='\t')

  # dataFilteredOut.to_csv(path_or_buf=OUTPUT_FILE_PATH +
'{}curated_out_peptides.txt'.format(CURRENT_TYPE), index=False, sep='\t')


"""PIPELINE : CONCATENATION OF RESULTS"""
#calls the function to concat and clean the df
finalTable = concatAndClean(allFormattedData)
# counts the AA occurences of all peptides and saves it into a specific file
countFinal = AA.countAAoccurence(finalTable)
countPercFinal = AA.computeAAocc(countFinal)
countPercFinal.to_csv(path_or_buf=OUTPUT_FILE_PATH + 'AAcounts_betaEnriched.txt',
                  index=False, sep='\t')
#saves if to a csv
finalTable.to_csv(path_or_buf=OUTPUT_FILE_PATH + 'CleanResultTable.txt', index=False, sep='\t')
```

References

1. Campbell-Thompson M, Fu A, Kaddis JS, et al. Insulitis and β-Cell Mass in the Natural History of Type 1 Diabetes. *Diabetes*. 2016;65(3):719-731. doi:10.2337/db15-0779

2. Pugliese A. Autoreactive T cells in type 1 diabetes. *J Clin Invest*. 2017;127(8):2881-2891. doi:10.1172/JCI94549

3. DiMeglio LA, Evans-Molina C, Oram RA. Type 1 diabetes. *Lancet*. 2018;391(10138):2449-2462. doi:10.1016/S0140-6736(18)31320-5

4. American Diabetes Association. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018. *Diabetes Care*. 2018;41(Supplement 1):S13-S27. doi:10.2337/dc18-S002

5. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157:107843. doi:10.1016/j.diabres.2019.107843

6. Mobasseri M, Shirmohammadi M, Amiri T, Vahed N, Hosseini Fard H, Ghojazadeh M. Prevalence and incidence of type 1 diabetes in the world: a systematic review and meta-analysis. *Health Promot Perspect*. 2020;10(2):98-115. doi:10.34172/hpp.2020.18

7. *IDF Diabetes Atlas 10th Edition*. International Diabetes Federation. Brussels, Belgium; 2021.

8. The Diamond Project Group. Incidence and trends of childhood Type 1 diabetes worldwide 1990–1999. *Diabetic Medicine*. 2006;23(8):857-866. doi:10.1111/j.1464-5491.2006.01925.x

9. Leete P, Mallone R, Richardson SJ, Sosenko JM, Redondo MJ, Evans-Molina C. The Effect of Age on the Progression and Severity of Type 1 Diabetes: Potential Effects on Disease Mechanisms. *Curr Diab Rep*. 2018;18(11):115. doi:10.1007/s11892-018-1083-4

10. Patterson CC, Karuranga S, Salpea P, et al. Worldwide estimates of incidence, prevalence and mortality of type 1 diabetes in children and adolescents: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. *Diabetes Res Clin Pract*. 2019;157:107842. doi:10.1016/j.diabres.2019.107842

11. Patterson CC, Harjutsalo V, Rosenbauer J, et al. Trends and cyclical variation in the incidence of childhood type 1 diabetes in 26 European centres in the 25 year period 1989-2013: a multicentre prospective registration study. *Diabetologia*. 2019;62(3):408-417. doi:10.1007/s00125-018-4763-3

12. Dabelea D, Mayer-Davis EJ, Saydah S, et al. Prevalence of Type 1 and Type 2 Diabetes Among Children and Adolescents From 2001 to 2009. *JAMA*. 2014;311(17):1778-1786. doi:10.1001/jama.2014.3201

13. Imperatore G, Boyle JP, Thompson TJ, et al. Projections of type 1 and type 2 diabetes burden in the U.S. population aged <20 years through 2050: dynamic modeling of

incidence, mortality, and population growth. *Diabetes Care*. 2012;35(12):2515-2520. doi:10.2337/dc12-0669

14. Norris JM, Johnson RK, Stene LC. Type 1 diabetes-early life origins and changing epidemiology. *Lancet Diabetes Endocrinol*. 2020;8(3):226-238. doi:10.1016/S2213-8587(19)30412-7

15. Eisenbarth GS. Type I diabetes mellitus. A chronic autoimmune disease. *N Engl J Med*. 1986;314(21):1360-1368. doi:10.1056/NEJM198605223142106

16. Kondrashova A, Viskari H, Kulmala P, et al. Signs of β-Cell Autoimmunity in Nondiabetic Schoolchildren: A comparison between Russian Karelia with a low incidence of type 1 diabetes and Finland with a high incidence rate. *Diabetes Care*. 2007;30(1):95-100. doi:10.2337/dc06-0711

17. Peled A, Gordon B, Twig G, et al. Immigration to Israel during childhood is associated with diabetes at adolescence: a study of 2.7 million adolescents. *Diabetologia*. 2017;60(11):2226-2230. doi:10.1007/s00125-017-4399-8

18. Söderström U, Aman J, Hjern A. Being born in Sweden increases the risk for type 1 diabetes - a study of migration of children to Sweden as a natural experiment. *Acta Paediatr*. 2012;101(1):73-77. doi:10.1111/j.1651-2227.2011.02410.x

19. Orban T, Sosenko JM, Cuthbertson D, et al. Pancreatic Islet Autoantibodies as Predictors of Type 1 Diabetes in the Diabetes Prevention Trial–Type 1. *Diabetes Care*. 2009;32(12):2269-2274. doi:10.2337/dc09-0934

20. Insel RA, Dunne JL, Atkinson MA, et al. Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care*. 2015;38(10):1964-1974. doi:10.2337/dc15-1419

21. Jacobsen LM, Bocchino L, Evans-Molina C, et al. The risk of progression to type 1 diabetes is highly variable in individuals with multiple autoantibodies following screening. *Diabetologia*. 2020;63(3):588-596. doi:10.1007/s00125-019-05047-w

22. Leete P, Willcox A, Krogvold L, et al. Differential Insulitic Profiles Determine the Extent of β-Cell Destruction and the Age at Onset of Type 1 Diabetes. *Diabetes*. 2016;65(5):1362-1369. doi:10.2337/db15-1615

23. Battaglia M, Ahmed S, Anderson MS, et al. Introducing the Endotype Concept to Address the Challenge of Disease Heterogeneity in Type 1 Diabetes. *Diabetes Care*. 2020;43(1):5-12.

24. Carré A, Richardson SJ, Larger E, Mallone R. Presumption of guilt for T cells in type 1 diabetes: lead culprits or partners in crime depending on age of onset? *Diabetologia*. 2021;64(1):15-25. doi:10.1007/s00125-020-05298-y

25. Barnett AH, Eff C, Leslie RD, Pyke DA. Diabetes in identical twins. A study of 200 pairs. *Diabetologia*. 1981;20(2):87-93. doi:10.1007/BF00262007

26. Redondo MJ, Jeffrey J, Fain PR, Eisenbarth GS, Orban T. Concordance for Islet Autoimmunity among Monozygotic Twins. 2010;359(26):2849-2850. doi:10.1056/NEJMc0805398

27. Aly TA, Ide A, Jahromi MM, et al. Extreme genetic risk for type 1A diabetes. *Proc Natl Acad Sci USA*. 2006;103(38):14074-14079. doi:10.1073/pnas.0606349103

28. Noble JA, Valdes AM, Varney MD, et al. HLA Class I and Genetic Susceptibility to Type 1 Diabetes. *Diabetes*. 2010;59(11):2972-2979. doi:10.2337/db10-0699

29. Nejentsev S, Howson JMM, Walker NM, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature*. 2007;450(7171):887-892. doi:10.1038/nature06406

30. Noble JA. Immunogenetics of type 1 diabetes: A comprehensive review. *J Autoimmun*. 2015;64:101-112. doi:10.1016/j.jaut.2015.07.014

31. Pugliese A, Zeller M, Fernandez A, et al. The insulin gene is transcribed in the human thymus and transcription levels correlate with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat Genet*. 1997;15(3):293-297. doi:10.1038/ng0397-293

32. Chentoufi AA, Polychronakos C. Insulin Expression Levels in the Thymus Modulate Insulin-Specific Autoreactive T-Cell Tolerance: The Mechanism by Which the IDDM2 Locus May Predispose to Diabetes. *Diabetes*. 2002;51(5):1383-1390. doi:10.2337/diabetes.51.5.1383

33. Bluestone JA, Herold K, Eisenbarth G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature*. 2010;464(7293):1293-1300. doi:10.1038/nature08933

34. Fløyel T, Kaur S, Pociot F. Genes Affecting β-Cell Function in Type 1 Diabetes. *Curr Diab Rep*. 2015;15(11):97. doi:10.1007/s11892-015-0655-9

35. Eizirik DL, Sammeth M, Bouckenooghe T, et al. The Human Pancreatic Islet Transcriptome: Expression of Candidate Genes for Type 1 Diabetes and the Impact of Pro-Inflammatory Cytokines. *PLOS Genetics*. 2012;8(3):e1002552. doi:10.1371/journal.pgen.1002552

36. Pociot F, Akolkar B, Concannon P, et al. Genetics of Type 1 Diabetes: What's Next? *Diabetes*. 2010;59(7):1561-1571. doi:10.2337/db10-0076

37. TEDDY Study Group. The Environmental Determinants of Diabetes in the Young (TEDDY) study: study design. *Pediatr Diabetes*. 2007;8(5):286-298. doi:10.1111/j.1399-5448.2007.00269.x

38. Craig ME, Nair S, Stein H, Rawlinson WD. Viruses and type 1 diabetes: a new look at an old story. *Pediatr Diabetes*. 2013;14(3):149-158. doi:10.1111/pedi.12033

39. Stene LC, Gale E a. M. The prenatal environment and type 1 diabetes. *Diabetologia*. 2013;56(9):1888-1897. doi:10.1007/s00125-013-2929-6

40. Rewers M, Ludvigsson J. Environmental risk factors for type 1 diabetes. *The Lancet*. 2016;387(10035):2340-2348. doi:10.1016/S0140-6736(16)30507-4

41. Gillespie KM, Bain SC, Barnett AH, et al. The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. *Lancet*. 2004;364(9446):1699-1700. doi:10.1016/S0140-6736(04)17357-1

42. Morgan NG, Richardson SJ. Enteroviruses as causative agents in type 1 diabetes: loose ends or lost cause? *Trends Endocrinol Metab*. 2014;25(12):611-619. doi:10.1016/j.tem.2014.08.002

43. Hyöty H. Viruses in type 1 diabetes. *Pediatric Diabetes*. 2016;17(S22):56-64. doi:10.1111/pedi.12370

44. Carré A, Vecchio F, Flodström-Tullberg M, You S, Mallone R. From association to causality: possible diabetogenic mechanisms of Coxsackievirus B infection and implications for type 1 diabetes prevention. *Endocr Rev*. submitted.

45. Ylipaasto P, Klingel K, Lindberg AM, et al. Enterovirus infection in human pancreatic islet cells, islet tropism in vivo and receptor involvement in cultured islet beta cells. *Diabetologia*. 2004;47(2):225-239. doi:10.1007/s00125-003-1297-z

46. Tracy S, Gauntt C. Group B Coxsackievirus Virulence. In: Tracy S, Oberste MS, Drescher KM, eds. Vol 323. Current Topics in Microbiology and Immunology. Springer Berlin Heidelberg; 2008:49-63. doi:10.1007/978-3-540-75546-3_3

47. Lönnrot M, Korpela K, Knip M, et al. Enterovirus Infection as a Risk Factor for β-Cell Autoimmunity in a Prospectively Observed Birth Cohort. 2000;49(8):1314-1318.

48. Laitinen OH, Honkanen H, Pakkanen O, et al. Coxsackievirus B1 Is Associated With Induction of β-Cell Autoimmunity That Portends Type 1 Diabetes. *Diabetes*. 2014;63(2):446-455. doi:10.2337/db13-0619

49. Vehik K, Lynch KF, Wong MC, et al. Prospective virome analyses in young children at increased genetic risk for type 1 diabetes. *Nat Med*. 2019;25(12):1865-1872. doi:10.1038/s41591-019-0667-0

50. Krogvold L, Edwin B, Buanes T, et al. Detection of a Low-Grade Enteroviral Infection in the Islets of Langerhans of Living Patients Newly Diagnosed With Type 1 Diabetes. *Diabetes*. 2015;64(5):1682-1687. doi:10.2337/db14-1370

51. Dotta F, Censini S, van Halteren AGS, et al. Coxsackie B4 virus infection of β cells and natural killer cell insulitis in recent-onset type 1 diabetic patients. *Proc Natl Acad Sci USA*. 2007;104(12):5115-5120. doi:10.1073/pnas.0700442104

52. Richardson SJ, Willcox A, Bone AJ, Foulis AK, Morgan NG. The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes. *Diabetologia*. 2009;52(6):1143-1151. doi:10.1007/s00125-009-1276-0

53. Atkinson MA, Bowman MA, Campbell L, Darrow BL, Kaufman DL, Maclaren NK. Cellular immunity to a determinant common to glutamate decarboxylase and coxsackie virus in insulin-dependent diabetes. *J Clin Invest*. 1994;94(5):2125-2129.

54. Stone VM, Hankaniemi MM, Svedin E, et al. A Coxsackievirus B vaccine protects against virus-induced diabetes in an experimental mouse model of type 1 diabetes. *Diabetologia*. 2018;61(2):476-481. doi:10.1007/s00125-017-4492-z

55. Stone VM, Hankaniemi MM, Laitinen OH, et al. A hexavalent Coxsackievirus B vaccine is highly immunogenic and has a strong protective capacity in mice and nonhuman primates. *Sci Adv*. 2020;6(19):eaaz2433. doi:10.1126/sciadv.aaz2433

56. Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. *N Engl J Med*. 2002;347(12):911-920. doi:10.1056/NEJMra020100

57. Pearson JA, Kakabadse D, Davies J, et al. Altered Gut Microbiota Activate and Expand Insulin B15-23–Reactive CD8+ T Cells. *Diabetes*. 2019;68(5):1002-1013. doi:10.2337/db18-0487

58. de Goffau MC, Luopajärvi K, Knip M, et al. Fecal microbiota composition differs between children with β-cell autoimmunity and those without. *Diabetes*. 2013;62(4):1238-1244. doi:10.2337/db12-0526

59. Kostic AD, Gevers D, Siljander H, et al. The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe*. 2015;17(2):260-273. doi:10.1016/j.chom.2015.01.001

60. Bosi E, Molteni L, Radaelli MG, et al. Increased intestinal permeability precedes clinical onset of type 1 diabetes. *Diabetologia*. 2006;49(12):2824-2827. doi:10.1007/s00125-006-0465-3

61. Sorini C, Cosorich I, Conte ML, et al. Loss of gut barrier integrity triggers activation of islet-reactive T cells and autoimmune diabetes. *Proc Natl Acad Sci USA*. 2019;116(30):15140-15149. doi:10.1073/pnas.1814558116

62. Girdhar K, Huang Q, Chow IT, et al. A gut microbial peptide and molecular mimicry in the pathogenesis of type 1 diabetes. *Proc Natl Acad Sci USA*. 2022;119(31):e2120028119. doi:10.1073/pnas.2120028119

63. Bach JF, Chatenoud L. The Hygiene Hypothesis: An Explanation for the Increased Frequency of Insulin-Dependent Diabetes. *Cold Spring Harb Perspect Med*. 2012;2(2):a007799. doi:10.1101/cshperspect.a007799

64. Gepts W. Pathologic Anatomy of the Pancreas in Juvenile Diabetes Mellitus. *Diabetes*. 1965;14(10):619-633. doi:10.2337/diab.14.10.619

65. Damond N, Engler S, Zanotelli VRT, et al. A Map of Human Type 1 Diabetes Progression by Imaging Mass Cytometry. *Cell Metab*. 2019;29(3):755-768.e5. doi:10.1016/j.cmet.2018.11.014

66. Zakharov PN, Hu H, Wan X, Unanue ER. Single-cell RNA sequencing of murine islets shows high cellular complexity at all stages of autoimmune diabetes. *J Exp Med*. 2020;217(6). doi:10.1084/jem.20192362

67. Serhan CN. Pro-resolving lipid mediators are leads for resolution physiology. *Nature*. 2014;510(7503):92-101. doi:10.1038/nature13479

68. Willcox A, Richardson SJ, Bone AJ, Foulis AK, Morgan NG. Analysis of islet inflammation in human type 1 diabetes. *Clin Exp Immunol*. 2009;155(2):173-181. doi:10.1111/j.1365-2249.2008.03860.x

69. Vecchio F, Lo Buono N, Stabilini A, et al. Abnormal neutrophil signature in the blood and pancreas of presymptomatic and symptomatic type 1 diabetes. *JCI Insight*. 2019;3(18). doi:10.1172/jci.insight.122146

70. Diana J, Simoni Y, Furio L, et al. Crosstalk between neutrophils, B-1a cells and plasmacytoid dendritic cells initiates autoimmune diabetes. *Nat Med*. 2013;19(1):65-73. doi:10.1038/nm.3042

71. Thayer TC, Delano M, Liu C, et al. Superoxide production by macrophages and T cells is critical for the induction of autoreactivity and type 1 diabetes. *Diabetes*. 2011;60(8):2144-2151. doi:10.2337/db10-1222

72. Zirpel H, Roep BO. Islet-Resident Dendritic Cells and Macrophages in Type 1 Diabetes: In Search of Bigfoot's Print. *Front Endocrinol*. 2021;12:666795. doi:10.3389/fendo.2021.666795

73. Calderon B, Suri A, Miller MJ, Unanue ER. Dendritic cells in islets of Langerhans constitutively present β cell-derived peptides bound to their class II MHC molecules. *Proc Natl Acad Sci USA*. 2008;105(16):6121-6126. doi:10.1073/pnas.0801973105

74. Rodriguez-Calvo T, Ekwall O, Amirian N, Zapardiel-Gonzalo J, von Herrath MG. Increased Immune Cell Infiltration of the Exocrine Pancreas: A Possible Contribution to the Pathogenesis of Type 1 Diabetes. *Diabetes*. 2014;63(11):3880-3890. doi:10.2337/db14-0549

75. Silveira PA, Grey ST. B cells in the spotlight: innocent bystanders or major players in the pathogenesis of type 1 diabetes. *Trends Endocrinol Metab*. 2006;17(4):128-135. doi:10.1016/j.tem.2006.03.006

76. Mariño E, Villanueva J, Walters S, Liuwantara D, Mackay F, Grey ST. CD4+CD25+ T-Cells Control Autoimmunity in the Absence of B-Cells. *Diabetes*. 2009;58(7):1568-1577. doi:10.2337/db08-1504

77. Wang YJ, Traum D, Schug J, et al. Multiplexed In Situ Imaging Mass Cytometry Analysis of the Human Endocrine Pancreas and Immune System in Type 1 Diabetes. *Cell Metab*. 2019;29(3):769-783.e4. doi:10.1016/j.cmet.2019.01.003

78. Wicker LS, Leiter EH, Todd JA, et al. β2-Microglobulin-Deficient NOD Mice Do Not Develop Insulitis or Diabetes. *Diabetes*. 1994;43(3):500-504. doi:10.2337/diab.43.3.500

79. Graham KL, Krishnamurthy B, Fynch S, et al. Autoreactive cytotoxic T lymphocytes acquire higher expression of cytotoxic effector markers in the islets of NOD mice after priming in pancreatic lymph nodes. *Am J Pathol*. 2011;178(6):2716-2725. doi:10.1016/j.ajpath.2011.02.015

80. Makhlouf L, Grey ST, Dong V, et al. Depleting anti-CD4 monoclonal antibody cures new-onset diabetes, prevents recurrent autoimmune diabetes, and delays allograft rejection in

nonobese diabetic mice. *Transplantation*. 2004;77(7):990-997. doi:10.1097/01.TP.0000118410.61419.59

81. Haskins K, McDuffie M. Acceleration of diabetes in young NOD mice with a CD4+ islet-specific T cell clone. *Science*. 1990;249(4975):1433-1436. doi:10.1126/science.2205920

82. Krogvold L, Wiberg A, Edwin B, et al. Insulitis and characterisation of infiltrating T cells in surgical pancreatic tail resections from patients at onset of type 1 diabetes. *Diabetologia*. 2016;59(3):492-501. doi:10.1007/s00125-015-3820-4

83. Tang Q, Adams JY, Penaranda C, et al. Central Role of Defective Interleukin-2 Production in the Triggering of Islet Autoimmune Destruction. *Immunity*. 2008;28(5):687-697. doi:10.1016/j.immuni.2008.03.016

84. Faucher P, Beuvon F, Fignani D, et al. Immunoregulated insulitis and slow-progressing type 1 diabetes after duodenopancreatectomy. *Diabetologia*. 2021;64(12):2731-2740. doi:10.1007/s00125-021-05563-8

85. Campbell-Thompson ML, Atkinson MA, Butler AE, et al. The diagnosis of insulitis in human type 1 diabetes. *Diabetologia*. 2013;56(11):2541-2543. doi:10.1007/s00125-013-3043-5

86. Ferreira RC, Guo H, Coulson RMR, et al. A Type I Interferon Transcriptional Signature Precedes Autoimmunity in Children Genetically at Risk for Type 1 Diabetes. *Diabetes*. 2014;63(7):2538-2550. doi:10.2337/db13-1777

87. Huang X, Yuan J, Goddard A, et al. Interferon Expression in the Pancreases of Patients With Type I Diabetes. *Diabetes*. 1995;44(6):658-664.

88. Li Q, Xu B, Michie SA, Rubins KH, Schreriber RD, McDevitt HO. Interferon-α initiates type 1 diabetes in nonobese diabetic mice. *Proc Natl Acad Sci USA*. 2008;105(34):12439-12444. doi:10.1073/pnas.0806439105

89. Lombardi A, Tsomos E, Hammerstad SS, Tomer Y. Interferon alpha: The key trigger of type 1 diabetes. *J Autoimmun*. 2018;94:7-15. doi:10.1016/j.jaut.2018.08.003

90. Campbell IL, Kay TW, Oxbrow L, Harrison LC. Essential role for interferon-gamma and interleukin-6 in autoimmune insulin-dependent diabetes in NOD/Wehi mice. *J Clin Invest*. 1991;87(2):739-742. doi:10.1172/JCI115055

91. von Herrath MG, Oldstone MB. Interferon-gamma is essential for destruction of beta cells and development of insulin-dependent diabetes mellitus. *J Exp Med*. 1997;185(3):531-539. doi:10.1084/jem.185.3.531

92. Savinov AY, Wong FS, Chervonsky AV. IFN-γ Affects Homing of Diabetogenic T Cells. *J Immunol*. 2001;167(11):6637-6643. doi:10.4049/jimmunol.167.11.6637

93. Suk K, Kim S, Kim YH, et al. IFN-γ/TNF-α Synergism as the Final Effector in Autoimmune Diabetes: A Key Role for STAT1/IFN Regulatory Factor-1 Pathway in Pancreatic β Cell Death. *J Immunol*. 2001;166(7):4481-4489. doi:10.4049/jimmunol.166.7.4481

94. Quattrin T, Haller MJ, Steck AK, et al. Golimumab and Beta-Cell Function in Youth with New-Onset Type 1 Diabetes. *N Engl J Med*. 2020;383(21):2007-2017. doi:10.1056/NEJMoa2006136

95. Lee MS, Wogensen L, Shizuru J, Oldstone MB, Sarvetnick N. Pancreatic islet production of murine interleukin-10 does not inhibit immune-mediated tissue destruction. *J Clin Invest*. 1994;93(3):1332-1338. doi:10.1172/JCI117092

96. Wogensen L, Lee MS, Sarvetnick N. Production of interleukin 10 by islet cells accelerates immune-mediated destruction of beta cells in nonobese diabetic mice. *J Exp Med*. 1994;179(4):1379-1384. doi:10.1084/jem.179.4.1379

97. Pakala SV, Kurrer MO, Katz JD. T helper 2 (Th2) T cells induce acute pancreatitis and diabetes in immune-compromised nonobese diabetic (NOD) mice. *J Exp Med*. 1997;186(2):299-306. doi:10.1084/jem.186.2.299

98. Robert S, Gysemans C, Takiishi T, et al. Oral delivery of glutamic acid decarboxylase (GAD)-65 and IL10 by Lactococcus lactis reverses diabetes in recent-onset NOD mice. *Diabetes*. 2014;63(8):2876-2887. doi:10.2337/db13-1236

99. Li C, Zhang L, Chen Y, Lin X, Li T. Protective role of adenovirus vector-mediated interleukin-10 gene therapy on endogenous islet β-cells in recent-onset type 1 diabetes in NOD mice. *Exp Ther Med*. 2016;11(5):1625-1632. doi:10.3892/etm.2016.3169

100. Kukreja A, Cost G, Marker J, et al. Multiple immuno-regulatory defects in type-1 diabetes. *J Clin Invest*. 2002;109(1):131-140. doi:10.1172/JCI13605

101. Kent SC, Chen Y, Clemmings SM, et al. Loss of IL-4 Secretion from Human Type 1a Diabetic Pancreatic Draining Lymph Node NKT Cells. *J Immunol*. 2005;175(7):4458-4464. doi:10.4049/jimmunol.175.7.4458

102. Wang B, Gonzalez A, Höglund P, Katz JD, Benoist C, Mathis D. Interleukin-4 Deficiency Does Not Exacerbate Disease in NOD Mice. *Diabetes*. 1998;47(8):1207-1211. doi:10.2337/diab.47.8.1207

103. Lu J, Liu J, Li L, Lan Y, Liang Y. Cytokines in type 1 diabetes: mechanisms of action and immunotherapeutic targets. *Clin Transl Immunology*. 2020;9(3):e1122. doi:10.1002/cti2.1122

104. Bingley PJ. Clinical Applications of Diabetes Antibody Testing. *J Clin Endocr Metab*. 2010;95(1):25-33. doi:10.1210/jc.2009-1365

105. Gottlieb PA, Eisenbarth GS. Diagnosis and treatment of pre-insulin dependent diabetes. *Annu Rev Med*. 1998;49:391-405. doi:10.1146/annurev.med.49.1.391

106. Palmer JP, Asplin CM, Clemons P, et al. Insulin Antibodies in Insulin-Dependent Diabetics Before Insulin Treatment. *Science*. 1983;222(4630):1337-1339. doi:10.1126/science.6362005

107. Baekkeskov S, Aanstoot HJ, Christgai S, et al. Identification of the 64K autoantigen in insulin-dependent diabetes as the GABA-synthesizing enzyme glutamic acid decarboxylase. *Nature*. 1990;347(6289):151-156. doi:10.1038/347151a0

108. Payton MA, Hawkes CJ, Christie MR. Relationship of the 37,000- and 40,000-M(r) tryptic fragments of islet antigens in insulin-dependent diabetes to the protein tyrosine phosphatase-like molecule IA-2 (ICA512). *J Clin Invest*. 1995;96(3):1506-1511. doi:10.1172/JCI118188

109. Wenzlau JM, Juhl K, Yu L, et al. The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes. *Proc Natl Acad Sci USA*. 2007;104(43):17040-17045. doi:10.1073/pnas.0705894104

110. Parikka V, Näntö-Salonen K, Saarinen M, et al. Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk. *Diabetologia*. 2012;55(7):1926-1936. doi:10.1007/s00125-012-2523-3

111. Ziegler AG, Rewers M, Simell O, et al. Seroconversion to Multiple Islet Autoantibodies and Risk of Progression to Diabetes in Children. *JAMA*. 2013;309(23):2473. doi:10.1001/jama.2013.6285

112. Endesfelder D, zu Castell W, Bonifacio E, et al. Time-Resolved Autoantibody Profiling Facilitates Stratification of Preclinical Type 1 Diabetes in Children. *Diabetes*. 2019;68(1):119-130. doi:10.2337/db18-0594

113. Patel SK, Ma CS, Fourlanos S, Greenfield JR. Autoantibody-Negative Type 1 Diabetes: A Neglected Subtype. *Trends Endocrinol Metab*. 2021;32(5):295-305. doi:10.1016/j.tem.2021.02.001

114. Deligne C, You S, Mallone R. Personalized Immunotherapies for Type 1 Diabetes: Who, What, When, and How? *J Pers Med*. 2022;12(4):542. doi:10.3390/jpm12040542

115. Mallone R, Roep BO. Biomarkers for immune intervention trials in type 1 diabetes. *Clin Immunol*. 2013;149(3):286-296. doi:10.1016/j.clim.2013.02.009

116. Christianson SW, Shultz LD, Leiter EH. Adoptive Transfer of Diabetes Into Immunodeficient NOD-scid/scid Mice: Relative Contributions of CD4+ and CD8+ T-Cells From Diabetic Versus Prediabetic NOD.NON-Thy-1a Donors. *Diabetes*. 1993;42(1):44-55. doi:10.2337/diab.42.1.44

117. Herold KC, Bundy BN, Long SA, et al. An Anti-CD3 Antibody, Teplizumab, in Relatives at Risk for Type 1 Diabetes. *N Engl J Med*. 2019;381(7):603-613. doi:10.1056/NEJMoa1902226

118. Gaglia J, Kissler S. Anti-CD3 Antibody for the Prevention of Type 1 Diabetes: A Story of Perseverance. *Biochemistry*. 2019;58(40):4107-4111. doi:10.1021/acs.biochem.9b00707

119. Culina S, Lalanne AI, Afonso G, et al. Islet-reactive CD8+ T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Sci Immunol*. 2018;3(20):eaao4013. doi:10.1126/sciimmunol.aao4013

120. Davis MM. Not-So-Negative Selection. *Immunity*. 2015;43(5):833-835. doi:10.1016/j.immuni.2015.11.002

121. Wing JB, Tanaka A, Sakaguchi S. Human FOXP3+ Regulatory T Cell Heterogeneity and Function in Autoimmunity and Cancer. *Immunity*. 2019;50(2):302-316. doi:10.1016/j.immuni.2019.01.020

122. Glisic-Milosavljevic S, Waukau J, Jailwala P, et al. At-Risk and Recent-Onset Type 1 Diabetic Subjects Have Increased Apoptosis in the CD4+CD25+high T-Cell Fraction. Sollid L, ed. *PLoS ONE*. 2007;2(1):e146. doi:10.1371/journal.pone.0000146

123. Long SA, Cerosaletti K, Bollyky PL, et al. Defects in IL-2R Signaling Contribute to Diminished Maintenance of FOXP3 Expression in CD4+CD25+ Regulatory T-Cells of Type 1 Diabetic Subjects. *Diabetes*. 2010;59(2):407-415. doi:10.2337/db09-0694

124. Leete P, McDonald TJ, Shields BM, et al. Studies of insulin and proinsulin in pancreas and serum support the existence of aetiopathological endotypes of type 1 diabetes associated with age at diagnosis. *Diabetologia*. 2020;63:1258-1267. doi:10.1007/s00125-020-05115-6

125. Pescovitz MD, Greenbaum CJ, Krause-Steinrauf H, et al. Rituximab, B-Lymphocyte Depletion, and Preservation of Beta-Cell Function. *N Engl J Med*. 2009;361(22):2143-2152. doi:10.1056/NEJMoa0904452

126. Sherry N, Hagopian W, Ludvigsson J, et al. Teplizumab for treatment of type 1 diabetes (Protégé study): 1-year results from a randomised, placebo-controlled trial. *Lancet*. 2011;378(9790):487-497. doi:10.1016/S0140-6736(11)60931-8

127. Hagopian W, Ferry RJ, Sherry N, et al. Teplizumab preserves C-peptide in recent-onset type 1 diabetes: two-year results from the randomized, placebo-controlled Protégé trial. *Diabetes*. 2013;62(11):3901-3908. doi:10.2337/db13-0236

128. Smith MJ, Cambier JC, Gottlieb PA. Endotypes in T1D: B lymphocytes and early onset. *Curr Opin Endocrinol Diabetes Obes*. 2020;27(4):225-230. doi:10.1097/MED.0000000000000547

129. Pociot F, Lernmark Å. Genetic risk factors for type 1 diabetes. *Lancet*. 2016;387(10035):2331-2339. doi:10.1016/S0140-6736(16)30582-7

130. Choo SY. The HLA System: Genetics, Immunology, Clinical Testing, and Clinical Implications. *Yonsei Med J*. 2007;48(1):11-23. doi:10.3349/ymj.2007.48.1.11

131. Klein J, Sato A. The HLA system. First of two parts. *N Engl J Med*. 2000;343(10):702-709. doi:10.1056/NEJM200009073431006

132. Yaneva R, Schneeweiss C, Zacharias M, Springer S. Peptide binding to MHC class I and II proteins: New avenues from new methods. *Mol Immunol*. 2010;47(4):649-657. doi:10.1016/j.molimm.2009.10.008

133. Wyatt RC, Lanzoni G, Russell MA, Gerling I, Richardson SJ. What the HLA-I!—Classical and Non-classical HLA Class I and Their Potential Roles in Type 1 Diabetes. *Curr Diab Rep*. 2019;19(12):159. doi:10.1007/s11892-019-1245-z

134. Braud VM, Allan DSJ, O'Callaghan CA, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature*. 1998;391(6669):795-799. doi:10.1038/35869

135. Joosten SA, Sullivan LC, Ottenhoff THM. Characteristics of HLA-E Restricted T-Cell Responses and Their Role in Infectious Diseases. *J Immunol Res*. 2016;2016:1-11. doi:10.1155/2016/2695396

136. Walters LC, Harlos K, Brackenridge S, et al. Pathogen-derived HLA-E bound epitopes reveal broad primary anchor pocket tolerability and conformationally malleable peptide binding. *Nat Commun*. 2018;9(1):3137. doi:10.1038/s41467-018-05459-z

137. Celik AA, Kraemer T, Huyton T, Blasczyk R, Bade-Döding C. The diversity of the HLA-E-restricted peptide repertoire explains the immunological impact of the Arg107Gly mismatch. *Immunogenetics*. 2016;68(1):29-41. doi:10.1007/s00251-015-0880-z

138. Coupel S, Moreau A, Hamidou M, Horejsi V, Soulillou JP, Charreau B. Expression and release of soluble HLA-E is an immunoregulatory feature of endothelial cell activation. *Blood*. 2007;109(7):2806-2814. doi:10.1182/blood-2006-06-030213

139. Boegel S, Löwer M, Bukur T, Sorn P, Castle JC, Sahin U. HLA and proteasome expression body map. *BMC Med Genomics*. 2018;11(1):36. doi:10.1186/s12920-018-0354-x

140. Hodgkinson AD, Millward BA, Demaine AG. The HLA-E locus is associated with age at onset and susceptibility to type 1 diabetes mellitus. *Hum Immunol*. 2000;61(3):290-295. doi:10.1016/S0198-8859(99)00116-0

141. Lin A, Yan WH. The Emerging Roles of Human Leukocyte Antigen-F in Immune Modulation and Viral Infection. *Front Immunol*. 2019;10. doi:10.3389/fimmu.2019.00964

142. Richardson SJ, Rodriguez-Calvo T, Gerling IC, et al. Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia*. 2016;59(11):2448-2458. doi:10.1007/s00125-016-4067-4

143. HoWangYin KY, Loustau M, Wu J, et al. Multimeric structures of HLA-G isoforms function through differential binding to LILRB receptors. *Cell Mol Life Sci*. 2012;69(23):4041-4049. doi:10.1007/s00018-012-1069-3

144. Stern LJ, Brown JH, Jardetzky TS, et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature*. 1994;368(6468):215-221. doi:10.1038/368215a0

145. Russell MA, Redick SD, Blodgett DM, et al. HLA Class II Antigen Processing and Presentation Pathway Components Demonstrated by Transcriptome and Protein Analyses of Islet β-Cells From Donors With Type 1 Diabetes. *Diabetes*. 2019;68(5):988-1001. doi:10.2337/db18-0686

146. Quesada-Masachs E, Zilberman S, Rajendran S, et al. Upregulation of HLA class II in pancreatic beta cells from organ donors with type 1 diabetes. *Diabetologia*. 2022;65(2):387-401. doi:10.1007/s00125-021-05619-9

147. Ackerman AL, Cresswell P. Cellular mechanisms governing cross-presentation of exogenous antigens. *Nat Immunol*. 2004;5(7):678-684. doi:10.1038/ni1082

148. Guermonprez P, Amigorena S. Pathways for antigen cross presentation. *Springer Semin Immunopathol*. 2005;26(3):257-271. doi:10.1007/s00281-004-0176-0

149. Dengjel J, Schoor O, Fischer R, et al. Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. *Proc Natl Acad Sci USA*. 2005;102(22):7922-7927. doi:10.1073/pnas.0501190102

150. Hershko A, Ciechanover A. The ubiquitin system. *Annu Rev Biochem*. 1998;67:425-479. doi:10.1146/annurev.biochem.67.1.425

151. Kisselev AF, Akopian TN, Woo KM, Goldberg AL. The sizes of peptides generated from protein by mammalian 26 and 20 S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *J Biol Chem*. 1999;274(6):3363-3371. doi:10.1074/jbc.274.6.3363

152. Cascio P, Hilton C, Kisselev AF, Rock KL, Goldberg AL. 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. *EMBO J*. 2001;20(10):2357-2366. doi:10.1093/emboj/20.10.2357

153. Neefjes J, Ovaa H. A peptide's perspective on antigen presentation to the immune system. *Nat Chem Biol*. 2013;9(12):769-775. doi:10.1038/nchembio.1391

154. Hulpke S, Tampé R. The MHC I loading complex: a multitasking machinery in adaptive immunity. *Trends Biochem Sci*. 2013;38(8):412-420. doi:10.1016/j.tibs.2013.06.003

155. Blum JS, Wearsch PA, Cresswell P. Pathways of Antigen Processing. *Annu Rev Immunol*. 2013;31:443-473. doi:10.1146/annurev-immunol-032712-095910

156. Nagarajan NA, de Verteuil DA, Sriranganadane D, et al. ERAAP Shapes the Peptidome Associated with Classical and Nonclassical MHC Class I Molecules. *J Immunol*. 2016;197(4):1035-1043. doi:10.4049/jimmunol.1500654

157. Medina F, Ramos M, Iborra S, de León P, Rodríguez-Castro M, Del Val M. Furin-processed antigens targeted to the secretory route elicit functional TAP1-/-CD8+ T lymphocytes in vivo. *J Immunol*. 2009;183(7):4639-4647. doi:10.4049/jimmunol.0901356

158. Oliveira CC, van Hall T. Alternative Antigen Processing for MHC Class I: Multiple Roads Lead to Rome. *Front Immunol*. 2015;6:298. doi:10.3389/fimmu.2015.00298

159. Vigneron N, Ferrari V, Van den Eynde BJ, Cresswell P, Leonhardt RM. Cytosolic Processing Governs TAP-Independent Presentation of a Critical Melanoma Antigen. *J Immunol*. 2018;201(7):1875-1888. doi:10.4049/jimmunol.1701479

160. Parmentier N, Stroobant V, Colau D, et al. Production of an antigenic peptide by insulin-degrading enzyme. *Nat Immunol*. 2010;11(5):449-454. doi:10.1038/ni.1862

161. York IA, Bhutani N, Zendzian S, Goldberg AL, Rock KL. Tripeptidyl peptidase II is the major peptidase needed to trim long antigenic precursors, but is not required for most MHC class I antigen presentation. *J Immunol*. 2006;177(3):1434-1443. doi:10.4049/jimmunol.177.3.1434

162. López D, Jiménez M, García-Calvo M, Del Val M. Concerted Antigen Processing of a Short Viral Antigen by Human Caspase-5 and -10. *J Biol Chem*. 2011;286(19):16910-16913. doi:10.1074/jbc.M111.234658

163. Neijssen J, Herberts C, Drijfhout JW, Reits E, Janssen L, Neefjes J. Cross-presentation by intercellular peptide transfer through gap junctions. *Nature*. 2005;434(7029):83-88. doi:10.1038/nature03290

164. Pang B, Neijssen J, Qiao X, et al. Direct Antigen Presentation and Gap Junction Mediated Cross-Presentation during Apoptosis. *J Immunol*. 2009;183(2):1083-1090. doi:10.4049/jimmunol.0900861

165. Livneh I, Cohen-Kaplan V, Cohen-Rosenzweig C, Avni N, Ciechanover A. The life cycle of the 26S proteasome: from birth, through regulation and function, and onto its death. *Cell Res*. 2016;26(8):869-885. doi:10.1038/cr.2016.86

166. McCarthy MK, Weinberg JB. The immunoproteasome and viral infection: a complex regulator of inflammation. *Front Microbiol*. 2015;6:21. doi:10.3389/fmicb.2015.00021

167. Murata S, Takahama Y, Kasahara M, Tanaka K. The immunoproteasome and thymoproteasome: functions, evolution and human disease. *Nat Immunol*. 2018;19(9):923-931. doi:10.1038/s41590-018-0186-z

168. Toes REM, Nussbaum AK, Degermann S, et al. Discrete Cleavage Motifs of Constitutive and Immunoproteasomes Revealed by Quantitative Analysis of Cleavage Products. *J Exp Med*. 2001;194(1):1-12.

169. Diez-Rivero CM, Lafuente EM, Reche PA. Computational analysis and modeling of cleavage by the immunoproteasome and the constitutive proteasome. *BMC Bioinformatics*. 2010;11:479. doi:10.1186/1471-2105-11-479

170. Dalet A, Stroobant V, Vigneron N, Van den Eynde BJ. Differences in the production of spliced antigenic peptides by the standard proteasome and the immunoproteasome. *Eur J Immunol*. 2011;41(1):39-46. doi:10.1002/eji.201040750

171. Mishto M, Liepe J, Textoris-Taube K, et al. Proteasome isoforms exhibit only quantitative differences in cleavage and epitope generation. *Eur J Immunol*. 2014;44(12):3508-3521. doi:10.1002/eji.201444902

172. Winter MB, La Greca F, Arastu-Kapur S, et al. Immunoproteasome functions explained by divergence in cleavage specificity and regulation. Marletta MA, ed. *eLife*. 2017;6:e27364. doi:10.7554/eLife.27364

173. Joffre OP, Segura E, Savina A, Amigorena S. Cross-presentation by dendritic cells. *Nat Rev Immunol*. 2012;12(8):557-569. doi:10.1038/nri3254

174. Embgenbroich M, Burgdorf S. Current Concepts of Antigen Cross-Presentation. *Front Immunol*. 2018;9(1643). doi:10.3389/fimmu.2018.01643

175. Montealegre S, van Endert PM. Endocytic Recycling of MHC Class I Molecules in Non-professional Antigen Presenting and Dendritic Cells. *Front Immunol*. 2019;9:3098. doi:10.3389/fimmu.2018.03098

176. Howarth M, Williams A, Tolstrup AB, Elliott T. Tapasin enhances MHC class I peptide presentation according to peptide half-life. *Proc Natl Acad Sci USA*. 2004;101(32):11737-11742. doi:10.1073/pnas.0306294101

177. Radhakrishna H, Donaldson JG. ADP-ribosylation factor 6 regulates a novel plasma membrane recycling pathway. *J Cell Biol*. 1997;139(1):49-61. doi:10.1083/jcb.139.1.49

178. Carré A, Mallone R. Making Insulin and Staying Out of Autoimmune Trouble: The Beta-Cell Conundrum. *Front Immunol*. 2021;12:639682. doi:10.3389/fimmu.2021.639682

179. Roche PA, Furuta K. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nat Rev Immunol*. 2015;15(4):203-216. doi:10.1038/nri3818

180. Abul K. A, Andrew H. L, Shiv P. *Cellular and Molecular Immunology*. 9th ed. Elsevier; 2018.

181. Fallas JL, Tobin HM, Lou O, Guo D, Sant'Angelo DB, Denzin LK. Ectopic expression of HLA-DO in mouse dendritic cells diminishes MHC class II antigen presentation. *J Immunol*. 2004;173(3):1549-1560. doi:10.4049/jimmunol.173.3.1549

182. Yi W, Seth NP, Martillotti T, Wucherpfennig KW, Sant'Angelo DB, Denzin LK. Targeted regulation of self-peptide presentation prevents type I diabetes in mice without disrupting general immunocompetence. *J Clin Invest*. 2010;120(4):1324-1336. doi:10.1172/JCI40220

183. Crotzer VL, Blum JS. Autophagy and Its Role in MHC-Mediated Antigen Presentation. *J Immunol*. 2009;182(6):3335-3341. doi:10.4049/jimmunol.0803458

184. Münz C. Autophagy Beyond Intracellular MHC Class II Antigen Presentation. *Trends Immunol*. 2016;37(11):755-763. doi:10.1016/j.it.2016.08.017

185. Adamopoulou E, Tenzer S, Hillen N, et al. Exploring the MHC-peptide matrix of central tolerance in the human thymus. *Nat Commun*. 2013;4(1):2039. doi:10.1038/ncomms3039

186. Aichinger G, Lechler RI. Endogenous pathway of class II presentation. *Biochem Soc Trans*. 1995;23(3):657-660. doi:10.1042/bst0230657

187. Leung CSK. Endogenous Antigen Presentation of MHC Class II Epitopes through Non-Autophagic Pathways. *Front Immunol*. 2015;6:464. doi:10.3389/fimmu.2015.00464

188. Gonzalez-Duque S, Azoury ME, Colli ML, et al. Conventional and Neo-antigenic Peptides Presented by β Cells Are Targeted by Circulating Naïve CD8+ T Cells in Type 1 Diabetic and Healthy Donors. *Cell Metab*. 2018;28(6):946-960. doi:10.1016/j.cmet.2018.07.007

189. Rajaraman S, Canjuga D, Ghosh M, et al. Measles Virus-Based Treatments Trigger a Pro-inflammatory Cascade and a Distinctive Immunopeptidome in Glioblastoma. *Mol Ther Oncolytics*. 2019;12:147-161. doi:10.1016/j.omto.2018.12.010

190. Kubiniok P, Marcu A, Bichmann L, et al. Understanding the constitutive presentation of MHC class I immunopeptidomes in primary tissues. *iScience*. 2022;25(2):103768. doi:10.1016/j.isci.2022.103768

191. Rammensee H, Bachmann J, Emmerich NP, Bachor OA, Stevanović S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*. 1999;50(3-4):213-219. doi:10.1007/s002510050595

192. Vita R, Mahajan S, Overton JA, et al. The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res*. 2019;47(D1):D339-D343. doi:10.1093/nar/gky1006

193. Purcell AW, Ramarathinam SH, Ternette N. Mass spectrometry–based identification of MHC-bound peptides for immunopeptidomics. *Nat Protoc*. 2019;14(6):1687-1707. doi:10.1038/s41596-019-0133-y

194. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. *Nucleic Acids Res*. 2020;48(W1):W449-W454. doi:10.1093/nar/gkaa379

195. Rapin N, Hoof I, Lund O, Nielsen M. MHC motif viewer. *Immunogenetics*. 2008;60(12):759-765. doi:10.1007/s00251-008-0330-2

196. Ghosh M, Di Marco M, Stevanović S. Identification of MHC Ligands and Establishing MHC Class I Peptide Motifs. In: van Endert P, ed. *Antigen Processing: Methods and Protocols*. Methods in Molecular Biology. Springer; 2019:137-147. doi:10.1007/978-1-4939-9450-2_11

197. O'Donnell TJ, Rubinsteyn A, Bonsack M, Riemer AB, Laserson U, Hammerbacher J. MHCflurry: Open-Source Class I MHC Binding Affinity Prediction. *Cell Syst*. 2018;7(1):129-132.e4. doi:10.1016/j.cels.2018.05.014

198. Gfeller D, Guillaume P, Michaux J, et al. The Length Distribution and Multiple Specificity of Naturally Presented HLA-I Ligands. *J Immunol*. 2018;201(12):3705-3716. doi:10.4049/jimmunol.1800914

199. Racle J, Michaux J, Rockinger GA, et al. Robust prediction of HLA class II epitopes by deep motif deconvolution of immunopeptidomes. *Nat Biotechnol*. 2019;37(11):1283-1286. doi:10.1038/s41587-019-0289-6

200. Alvarez B, Reynisson B, Barra C, et al. NNAlign_MA; MHC Peptidome Deconvolution for Accurate MHC Binding Motif Characterization and Improved T-cell Epitope Predictions. *Mol Cell Proteomics*. 2019;18(12):2459-2477. doi:10.1074/mcp.TIR119.001658

201. Amengual-Rigo P, Guallar V. NetCleave: an open-source algorithm for predicting C-terminal antigen processing for MHC-I and MHC-II. *Sci Rep*. 2021;11(1):13126. doi:10.1038/s41598-021-92632-y

202. Nielsen M, Lund O, Buus S, Lundegaard C. MHC Class II epitope predictive algorithms. *Immunology*. 2010;130(3):319-328. doi:10.1111/j.1365-2567.2010.03268.x

203. Lanoix J, Durette C, Courcelles M, et al. Comparison of the MHC I Immunopeptidome Repertoire of B-Cell Lymphoblasts Using Two Isolation Methods. *Proteomics*. 2018;18(12):1700251. doi:10.1002/pmic.201700251

204. Mommen GPM, Frese CK, Meiring HD, et al. Expanding the detectable HLA peptide repertoire using electron-transfer/higher-energy collision dissociation (EThcD). *Proc Natl Acad Sci USA*. 2014;111(12):4507-4512. doi:10.1073/pnas.1321458111

205. Bassani-Sternberg M, Bräunlein E, Klar R, et al. Direct identification of clinically relevant neoepitopes presented on native human melanoma tissue by mass spectrometry. *Nat Commun*. 2016;7:13404. doi:10.1038/ncomms13404

206. Hu A, Noble WS, Wolf-Yadlin A. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Res*. 2016;5:419. doi:10.12688/f1000research.7042.1

207. Tabb DL, Vega-Montoto L, Rudnick PA, et al. Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res*. 2010;9(2):761-776. doi:10.1021/pr9006365

208. Gillet LC, Navarro P, Tate S, et al. Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis. *Mol Cell Proteomics*. 2012;11(6):O111.016717. doi:10.1074/mcp.O111.016717

209. Listgarten J, Emili A. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Mol Cell Proteomics*. 2005;4(4):419-434. doi:10.1074/mcp.R500005-MCP200

210. Matthiesen R, Jensen ON. Analysis of mass spectrometry data in proteomics. *Methods Mol Biol*. 2008;453:105-122. doi:10.1007/978-1-60327-429-6_4

211. Chen C, Hou J, Tanner JJ, Cheng J. Bioinformatics Methods for Mass Spectrometry-Based Proteomics Data Analysis. *Int J Mol Sci*. 2020;21(8):2873. doi:10.3390/ijms21082873

212. UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*. 2021;49(D1):D480-D489. doi:10.1093/nar/gkaa1100

213. Nesvizhskii AI. Proteogenomics: concepts, applications and computational strategies. *Nat Methods*. 2014;11(11):1114-1125. doi:10.1038/nmeth.3144

214. Kovalchik KA, Ma Q, Wessling L, et al. MhcVizPipe: A Quality Control Software for Rapid Assessment of Small- to Large-Scale Immunopeptidome Datasets. *Mol Cell Proteomics*. 2022;21(1):100178. doi:10.1016/j.mcpro.2021.100178

215. Lam H. Building and Searching Tandem Mass Spectral Libraries for Peptide Identification. *Mol Cell Proteomics*. 2011;10(12):R111.008565. doi:10.1074/mcp.R111.008565

216. Zhang F, Ge W, Ruan G, Cai X, Guo T. Data-Independent Acquisition Mass Spectrometry-Based Proteomics and Software Tools: A Glimpse in 2020. *Proteomics*. 2020;20(17-18):1900276. doi:10.1002/pmic.201900276

217. MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010;26(7):966-968. doi:10.1093/bioinformatics/btq054

218. Bruderer R, Bernhardt OM, Gandhi T, et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated

three-dimensional liver microtissues. *Mol Cell Proteomics*. 2015;14(5):1400-1410. doi:10.1074/mcp.M114.044305

219. Röst HL, Liu Y, D'Agostino G, et al. TRIC: an automated alignment strategy for reproducible protein quantification in targeted proteomics. *Nat Methods*. 2016;13(9):777-783. doi:10.1038/nmeth.3954

220. Bassani-Sternberg M, Digklia A, Huber F, et al. A Phase Ib Study of the Combination of Personalized Autologous Dendritic Cell Vaccine, Aspirin, and Standard of Care Adjuvant Chemotherapy Followed by Nivolumab for Resected Pancreatic Adenocarcinoma—A Proof of Antigen Discovery Feasibility in Three Patients. *Front Immunol*. 2019;10. Accessed June 21, 2022. https://www.frontiersin.org/article/10.3389/fimmu.2019.01832

221. Blass E, Ott PA. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. *Nat Rev Clin Oncol*. 2021;18(4):215-229. doi:10.1038/s41571-020-00460-2

222. Harari A, Sarivalasis A, de Jonge K, et al. A Personalized Neoantigen Vaccine in Combination with Platinum-Based Chemotherapy Induces a T-Cell Response Coinciding with a Complete Response in Endometrial Carcinoma. *Cancers*. 2021;13(22):5801. doi:10.3390/cancers13225801

223. Mohsen MO, Speiser DE, Michaux J, et al. Bedside formulation of a personalized multi-neoantigen vaccine against mammary carcinoma. *J Immunother Cancer*. 2022;10(1):e002927. doi:10.1136/jitc-2021-002927

224. Marko-Varga G, Fehniger TE. Proteomics and Disease The Challenges for Technology and Discovery. *J Proteome Res*. 2004;3(2):167-178. doi:10.1021/pr049958+

225. Kim MS, Zhong J, Pandey A. Common errors in mass spectrometry-based analysis of post-translational modifications. *Proteomics*. 2016;16(5):700-714. doi:10.1002/pmic.201500355

226. Mei S, Ayala R, Ramarathinam SH, et al. Immunopeptidomic Analysis Reveals That Deamidated HLA-bound Peptides Arise Predominantly from Deglycosylated Precursors. *Mol Cell Proteomics*. 2020;19(7):1236-1247. doi:10.1074/mcp.RA119.001846

227. Trigwell SM, Radford PM, Page SR, et al. Islet glutamic acid decarboxylase modified by reactive oxygen species is recognized by antibodies from patients with type 1 diabetes mellitus. *Clin Exp Immunol*. 2001;126(2):242-249. doi:10.1046/j.1365-2249.2001.01653.x

228. Khan MWA, Sherwani S, Khan WA, Moinuddin, Ali R. Characterization of hydroxyl radical modified GAD $_{65}$: A potential autoantigen in type 1 diabetes. *Autoimmunity*. 2009;42(2):150-158. doi:10.1080/08916930802468276

229. van Lummel M, Duinkerken G, van Veelen PA, et al. Posttranslational Modification of HLA-DQ Binding Islet Autoantigens in Type 1 Diabetes. *Diabetes*. 2014;63(1):237-247. doi:10.2337/db12-1214

230. McLaughlin RJ, de Haan A, Zaldumbide A, et al. Human islets and dendritic cells generate post-translationally modified islet autoantigens. *Clin Exp Immunol*. 2016;185(2):133-140. doi:10.1111/cei.12775

231. Rondas D, Crèvecoeur I, D'Hertog W, et al. Citrullinated Glucose-Regulated Protein 78 Is an Autoantigen in Type 1 Diabetes. *Diabetes*. 2014;64(2):573-586. doi:10.2337/db14-0621

232. Azoury ME, Samassa F, Buitinga M, et al. CD8+ T Cells Variably Recognize Native Versus Citrullinated GRP78 Epitopes in Type 1 Diabetes. *Diabetes*. 2021;70(12):2879-2891. doi:10.2337/db21-0259

233. Mei-Ling Y, Sean E. C, Renelle G, et al. Carbonyl Post-Translational Modification Associated with Early Onset Type 1 Diabetes Autoimmunity. *Diabetes*. 2022;71(9):db210989. doi:10.2337/db21-0989

234. Liepe J, Marino F, Sidney J, et al. A large fraction of HLA class I ligands are proteasome-generated spliced peptides. *Science*. 2016;354(6310):354-358. doi:10.1126/science.aaf4384

235. Mylonas R, Beer I, Iseli C, et al. Estimating the Contribution of Proteasomal Spliced Peptides to the HLA-I Ligandome. *Mol Cell Proteomics*. 2018;17(12):2347-2357. doi:10.1074/mcp.RA118.000877

236. Faridi P, Li C, Ramarathinam SH, et al. A subset of HLA-I peptides are not genomically templated: Evidence for cis- and trans-spliced peptide ligands. *Sci Immunol*. 2018;3(28):eaar3947. doi:10.1126/sciimmunol.aar3947

237. Lichti CF, Vigneron N, Clauser KR, Van den Eynde BJ, Bassani-Sternberg M. Navigating Critical Challenges Associated with Immunopeptidomics-Based Detection of Proteasomal Spliced Peptide Candidates. *Cancer Immunol Res*. Published online February 1, 2022:canimm;2326-6066.CIR-21-0727v1. doi:10.1158/2326-6066.CIR-21-0727

238. Delong T, Wiles TA, Baker RL, et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science*. 2016;351(6274):711-714. doi:10.1126/science.aad2791

239. Aaron Wiles T, Powell R, Michel CR, et al. Identification of Hybrid Insulin Peptides (HIPs) in Mouse and Human Islets by Mass Spectrometry. *J Proteome Res*. 2019;18(3):814-825. doi:10.1021/acs.jproteome.8b00875

240. Laumont CM, Vincent K, Hesnard L, et al. Noncoding regions are the main source of targetable tumor-specific antigens. *Sci Transl Med*. 2018;10(eaau5516).

241. Rolfs Z, Solntsev SK, Shortreed MR, Frey BL, Smith LM. Global Identification of Post-Translationally Spliced Peptides with Neo-Fusion. *J Proteome Res*. 2018;18:349-358. doi:10.1021/acs.jproteome.8b00651

242. Wilhelm M, Zolg DP, Graber M, et al. Deep learning boosts sensitivity of mass spectrometry-based immunopeptidomics. *Nat Commun*. 2021;12(1):3346. doi:10.1038/s41467-021-23713-9

243. Liao H, Barra C, Peng X, et al. *MARS: Improved De Novo Peptide Candidate Selection for Non-Canonical Antigen Target Discovery in Cancer*. In Review; 2022. doi:10.21203/rs.3.rs-1890352/v1, Preprint

244. Mishto M, Liepe J. Post-Translational Peptide Splicing and T Cell Responses. *Trends Immunol*. 2017;38(12):904-915. doi:10.1016/j.it.2017.07.011

245. James EA, Mallone R, Kent SC, DiLorenzo TP. T-Cell Epitopes and Neo-epitopes in Type 1 Diabetes: A Comprehensive Update and Reappraisal. *Diabetes*. 2020;69(7):1311-1335. doi:10.2337/dbi19-0022

246. Baker RL, Rihanek M, Hohenstein AC, et al. Hybrid Insulin Peptides Are Autoantigens in Type 1 Diabetes. *Diabetes*. 2019;68(9):1830-1840. doi:10.2337/db19-0128

247. Crawford SA, Wiles TA, Wenzlau JM, et al. Cathepsin D Drives the Formation of Hybrid Insulin Peptides Relevant to the Pathogenesis of Type 1 Diabetes. *Diabetes*. Published online August 30, 2022:db220303. doi:10.2337/db22-0303

248. Laumont CM, Daouda T, Laverdure JP, et al. Global proteogenomic analysis of human MHC class I-associated peptides derived from non-canonical reading frames. *Nat Commun*. 2016;7(1):10238. doi:10.1038/ncomms10238

249. Chen J, Brunner AD, Cogan JZ, et al. Pervasive functional translation of noncanonical human open reading frames. *Science*. 2020;367(6482):1140-1146. doi:10.1126/science.aay0262

250. Chong C, Müller M, Pak H, et al. Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. *Nat Commun*. 2020;11. doi:10.1038/s41467-020-14968-9

251. Smart AC, Margolis CA, Pimentel H, et al. Intron retention is a source of neoepitopes in cancer. *Nat Biotechnol*. 2018;36(11):1056-1058. doi:10.1038/nbt.4239

252. Ouspenskaia T, Law T, Clauser KR, et al. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat Biotechnol*. 2022;40(2):209-217. doi:10.1038/s41587-021-01021-3

253. Laumont CM, Perreault C. Exploiting non-canonical translation to identify new targets for T cell-based cancer immunotherapy. *Cell Mol Life Sci*. 2018;75(4):607-621. doi:10.1007/s00018-017-2628-4

254. Samassa F, Mallone R. Self-antigens, benign autoimmunity and type 1 diabetes: a beta-cell and T-cell perspective. *Curr Opin Endocrinol Diabetes Obes*. 2022;29(4):370-378. doi:10.1097/MED.0000000000000735

255. Phillips JM, Parish NM, Raine T, et al. Type 1 diabetes development requires both CD4+ and CD8+ T cells and can be reversed by non-depleting antibodies targeting both T cell populations. *Rev Diabet Stud*. 2009;6(2):97-103. doi:10.1900/RDS.2009.6.97

256. Campbell-Thompson M, Wasserfall C, Kaddis J, et al. Network for Pancreatic Organ Donors with Diabetes (nPOD): developing a tissue biobank for type 1 diabetes. *Diabetes Metab Res Rev*. 2012;28(7):608-617. doi:10.1002/dmrr.2316

257. Knabel M, Franz TJ, Schiemann M, et al. Reversible MHC multimer staining for functional isolation of T-cell populations and effective adoptive transfer. *Nat Med*. 2002;8(6):631-637. doi:10.1038/nm0602-631

258. Hadrup SR, Bakker AH, Shu CJ, et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods*. 2009;6(7):520-526. doi:10.1038/nmeth.1345

259. Andersen RS, Kvistborg P, Frøsig TM, et al. Parallel detection of antigen-specific T cell responses by combinatorial encoding of MHC multimers. *Nat Protoc*. 2012;7(5):891-902. doi:10.1038/nprot.2012.037

260. Azoury ME, Tarayrah M, Afonso G, et al. Peptides Derived From Insulin Granule Proteins Are Targeted by CD8+ T Cells Across MHC Class I Restrictions in Humans and NOD Mice. *Diabetes*. 2020;69(12):2678-2690. doi:10.2337/db20-0013

261. Yeo L, Pujol-Autonell I, Baptista R, et al. Circulating β cell-specific CD8 $^{+}$ T cells restricted by high-risk HLA class I molecules show antigen experience in children with and at risk of type 1 diabetes. *Clin Exp Immunol*. 2020;199(3):263-277. doi:10.1111/cei.13391

262. Skowera A, Ladell K, McLaren JE, et al. β-cell-specific CD8 T cell phenotype in type 1 diabetes reflects chronic autoantigen exposure. *Diabetes*. 2015;64(3):916-925. doi:10.2337/db14-0332

263. Reijonen H, Novak EJ, Kochik S, et al. Detection of GAD65-Specific T-Cells by Major Histocompatibility Complex Class II Tetramers in Type 1 Diabetic Patients and At-Risk Subjects. *Diabetes*. 2002;51(5):1375-1382. doi:10.2337/diabetes.51.5.1375

264. Öling V, Marttila J, Ilonen J, et al. GAD65- and proinsulin-specific CD4+ T-cells detected by MHC class II tetramers in peripheral blood of type 1 diabetes patients and at-risk subjects. *J Autoimmun*. 2005;25(3):235-243. doi:10.1016/j.jaut.2005.09.018

265. Gómez-Touriño I, Simón-Vázquez R, Alonso-Lorenzo J, et al. Characterization of the autoimmune response against the nerve tissue S100β in patients with type 1 diabetes. *Clin Exp Immunol*. 2015;180(2):207-217. doi:10.1111/cei.12572

266. Coppieters KT, Dotta F, Amirian N, et al. Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients. *J Exp Med*. 2012;209(1):51-60. doi:10.1084/jem.20111187

267. Calderon B, Carrero JA, Miller MJ, Unanue ER. Cellular and molecular events in the localization of diabetogenic T cells to islets of Langerhans. *Proc Natl Acad Sci USA*. 2011;108(4):1561-1566. doi:10.1073/pnas.1018973108

268. Rodriguez-Calvo T, Christoffersson G, Bender C, et al. Means, Motive, and Opportunity: Do Non-Islet-Reactive Infiltrating T Cells Contribute to Autoimmunity in Type 1 Diabetes? *Front Immunol*. 2021;12:683091. doi:10.3389/fimmu.2021.683091

269. Alkemade GM, Clemente-Casares X, Yu Z, et al. Local autoantigen expression as essential gatekeeper of memory T-cell recruitment to islet grafts in diabetic hosts. *Diabetes*. 2013;62(3):905-911. doi:10.2337/db12-0600

270. Burns J, Rosenzweig A, Zweiman B, Lisak RP. Isolation of myelin basic protein-reactive T-cell lines from normal human blood. *Cell Immunol*. 1983;81(2):435-440. doi:10.1016/0008-8749(83)90250-2

271. Lindert RB, Haase CG, Brehm U, Linington C, Wekerle H, Hohlfeld R. Multiple sclerosis: B- and T-cell responses to the extracellular domain of the myelin oligodendrocyte glycoprotein. *Brain*. 1999;122(11):2089-2100. doi:10.1093/brain/122.11.2089

272. Sabatino JJ, Wilson MR, Calabresi PA, Hauser SL, Schneck JP, Zamvil SS. Anti-CD20 therapy depletes activated myelin-specific CD8+ T cells in multiple sclerosis. *Proc Natl Acad Sci USA*. 2019;116(51):25800-25807. doi:10.1073/pnas.1915309116

273. Velthuis JH, Unger WW, Abreu JRF, et al. Simultaneous detection of circulating autoreactive CD8+ T-cells specific for different islet cell-associated epitopes using combinatorial MHC multimers. *Diabetes*. 2010;59(7):1721-1730. doi:10.2337/db09-1486

274. Viglietta V, Kent SC, Orban T, Hafler DA. GAD65-reactive T cells are activated in patients with autoimmune type 1a diabetes. *J Clin Invest*. 2002;109(7):895-903. doi:10.1172/JCI14114

275. Rius C, Attaf M, Tungatt K, et al. Peptide-MHC Class I Tetramers Can Fail To Detect Relevant Functional T Cell Clonotypes and Underestimate Antigen-Reactive T Cell Populations. *J Immunol*. 2018;200(7):2263-2279. doi:10.4049/jimmunol.1700242

276. Gattinoni L, Lugli E, Ji Y, et al. A human memory T cell subset with stem cell–like properties. *Nat Med*. 2011;17(10):1290-1297. doi:10.1038/nm.2446

277. van den Broek T, Borghans JAM, van Wijk F. The full spectrum of human naive T cells. *Nat Rev Immunol*. 2018;18(6):363-373. doi:10.1038/s41577-018-0001-y

278. Mallone R, Eizirik DL. Presumption of innocence for beta cells: why are they vulnerable autoimmune targets in type 1 diabetes? *Diabetologia*. 2020;63(10):1999-2006. doi:10.1007/s00125-020-05176-7

279. Mallone R, Halliez C, Rui J, Herold KC. The β-Cell in Type 1 Diabetes Pathogenesis: A Victim of Circumstances or an Instigator of Tragic Events? *Diabetes*. 2022;71(8):1603-1610. doi:10.2337/dbi21-0036

280. Eizirik DL, Miani M, Cardozo AK. Signalling danger: endoplasmic reticulum stress and the unfolded protein response in pancreatic islet inflammation. *Diabetologia*. 2013;56(2):234-241. doi:10.1007/s00125-012-2762-3

281. Oslowski CM, Hara T, O'Sullivan-Murphy B, et al. Thioredoxin-interacting protein mediates ER stress-induced β cell death through initiation of the inflammasome. *Cell Metab*. 2012;16(2):265-273. doi:10.1016/j.cmet.2012.07.005

282. Marré ML, James EA, Piganelli JD. β cell ER stress and the implications for immunogenicity in type 1 diabetes. *Front Cell Dev Biol*. 2015;3:67. doi:10.3389/fcell.2015.00067

283. van Kuppeveld FJ, Hoenderop JG, Smeets RL, et al. Coxsackievirus protein 2B modifies endoplasmic reticulum membrane and plasma membrane permeability and facilitates virus release. *EMBO J*. 1997;16(12):3519-3532. doi:10.1093/emboj/16.12.3519

284. Fung TS, Torres J, Liu DX. The Emerging Roles of Viroporins in ER Stress Response and Autophagy Induction during Virus Infection. *Viruses*. 2015;7(6):2834-2857. doi:10.3390/v7062749

285. Colli ML, Paula FM, Marselli L, et al. Coxsackievirus B Tailors the Unfolded Protein Response to Favour Viral Amplification in Pancreatic β Cells. *J Innate Immun*. 2019;11(4):375-390. doi:10.1159/000496034

286. Cornell CT, Kiosses WB, Harkins S, Whitton JL. Coxsackievirus B3 Proteins Directionally Complement Each Other To Downregulate Surface Major Histocompatibility Complex Class I. *J Virol*. 2007;81(13):6785-6797. doi:10.1128/JVI.00198-07

287. Knoch KP, Nath-Sain S, Petzold A, et al. PTBP1 is required for glucose-stimulated cap-independent translation of insulin granule proteins and Coxsackieviruses in beta cells. *Mol Metab*. 2014;3(5):518-530. doi:10.1016/j.molmet.2014.05.002

288. Kemball CC, Alirezaei M, Flynn CT, et al. Coxsackievirus Infection Induces Autophagy-Like Vesicles and Megaphagosomes in Pancreatic Acinar Cells In Vivo. *J Virol*. 2010;84(23):12110-12124. doi:10.1128/JVI.01417-10

289. Shi X, Chen Z, Tang S, Wu F, Xiong S, Dong C. Coxsackievirus B3 infection induces autophagic flux, and autophagosomes are critical for efficient viral replication. *Arch Virol*. 2016;161(8):2197-2205. doi:10.1007/s00705-016-2896-6

290. Mohamud Y, Shi J, Qu J, et al. Enteroviral Infection Inhibits Autophagic Flux via Disruption of the SNARE Complex to Enhance Viral Replication. *Cell Reports*. 2018;22(12):3292-3303. doi:10.1016/j.celrep.2018.02.090

291. Marroqui L, Masini M, Merino B, et al. Pancreatic α Cells are Resistant to Metabolic Stress-induced Apoptosis in Type 2 Diabetes. *EBioMedicine*. 2015;2(5):378-385. doi:10.1016/j.ebiom.2015.03.012

292. Op de Beeck A, Eizirik DL. Viral infections in type 1 diabetes mellitus — why the β cells? *Nat Rev Endocrinol*. 2016;12(5):263-273. doi:10.1038/nrendo.2016.30

293. Nigi L, Brusco N, Grieco GE, et al. Increased Expression of Viral Sensor MDA5 in Pancreatic Islets and in Hormone-Negative Endocrine Cells in Recent Onset Type 1 Diabetic Donors. *Front Immunol*. 2022;13(833141).

294. Chehadeh W, Kerr-Conte J, Pattou F, et al. Persistent Infection of Human Pancreatic Islets by Coxsackievirus B Is Associated with Alpha Interferon Synthesis in β Cells. *J Virol*. 2000;74(21):10153-10164.

295. Ylipaasto P, Smura T, Gopalacharyulu P, et al. Enterovirus-induced gene expression profile is critical for human pancreatic islet destruction. *Diabetologia*. 2012;55(12):3273-3283. doi:10.1007/s00125-012-2713-z

296. Kallionpää H, Elo LL, Laajala E, et al. Innate immune activity is detected prior to seroconversion in children with HLA-conferred type 1 diabetes susceptibility. *Diabetes*. 2014;63(7):2402-2414. doi:10.2337/db13-1775

297. Oshima M, Knoch KP, Diedisheim M, et al. Virus-like infection induces human β cell dedifferentiation. *JCI Insight*. 2018;3(3). doi:10.1172/jci.insight.97732

298. Marhfour I, Lopez XM, Lefkaditis D, et al. Expression of endoplasmic reticulum stress markers in the islets of patients with type 1 diabetes. *Diabetologia*. 2012;55(9):2417-2420. doi:10.1007/s00125-012-2604-3

299. Colli ML, Ramos-Rodríguez M, Nakayasu ES, et al. An integrated multi-omics approach identifies the landscape of interferon-α-mediated responses of human pancreatic beta cells. *Nat Commun*. 2020;11:2584. doi:10.1038/s41467-020-16327-0

300. Marré ML, Profozich JL, Coneybeer JT, et al. Inherent ER stress in pancreatic islet β cells causes self-recognition by autoreactive T cells in type 1 diabetes. *J Autoimmun*. 2016;72:33-46. doi:10.1016/j.jaut.2016.04.009

301. Marroqui L, Dos Santos RS, Op de Beeck A, et al. Interferon-α mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes. *Diabetologia*. 2017;60(4):656-667. doi:10.1007/s00125-016-4201-3

302. Coomans de Brachène A, Dos Santos RS, Marroqui L, et al. IFN-α induces a preferential long-lasting expression of MHC class I in human pancreatic beta cells. *Diabetologia*. 2018;61(3):636-640. doi:10.1007/s00125-017-4536-4

303. Babon JAB, DeNicola ME, Blodgett DM, et al. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. *Nat Med*. 2016;22(12):1482-1487. doi:10.1038/nm.4203

304. Naserke HE, Ziegler AG, Lampasona V, Bonifacio E. Early Development and Spreading of Autoantibodies to Epitopes of IA-2 and Their Association with Progression to Type 1 Diabetes. *J Immunol*. 1998;161(12):6963-6969.

305. Yu M, Johnson JM, Tuohy VK. A predictable sequential determinant spreading cascade invariably accompanies progression of experimental autoimmune encephalomyelitis: a basis for peptide-specific therapy after onset of clinical disease. *J Exp Med*. 1996;183(4):1777-1788. doi:10.1084/jem.183.4.1777

306. Krishnamurthy B, Dudek NL, McKenzie MD, et al. Responses against islet antigens in NOD mice are prevented by tolerance to proinsulin but not IGRP. *J Clin Invest*. 2006;116(12):3258-3265. doi:10.1172/JCI29602

307. Smeekens SP, Montag AG, Thomas G, et al. Proinsulin processing by the subtilisin-related proprotein convertases furin, PC2, and PC3. *Proc Natl Acad Sci USA*. 1992;89(18):8822-8826. doi:10.1073/pnas.89.18.8822

308. Liu M, Wright J, Guo H, Xiong Y, Arvan P. Chapter Two - Proinsulin Entry and Transit Through the Endoplasmic Reticulum in Pancreatic Beta Cells. In: Litwack G, ed. *Vitamins*

& Hormones. Vol 95. The Pancreatic Beta Cell. Academic Press; 2014:35-62. doi:10.1016/B978-0-12-800174-5.00002-8

309. Mbikay M, Seidah NG, Chrétien M. Neuroendocrine secretory protein 7B2: structure, expression and functions. *Biochem J*. 2001;357:329-342. doi:10.1042/0264-6021:3570329

310. Bartolomucci A, Possenti R, Mahata SK, Fischer-Colbrie R, Loh YP, Salton SRJ. The Extended Granin Family: Structure, Function, and Biomedical Implications. *Endocr Rev*. 2011;32(6):755-797. doi:10.1210/er.2010-0027

311. van der Meulen T, Donaldson CJ, Cáceres E, et al. Urocortin3 mediates somatostatin-dependent negative feedback control of insulin secretion. *Nat Med*. 2015;21(7):769-776. doi:10.1038/nm.3872

312. Roep BO, Arden SD, De Vries RRP, Hutton JC. T-cell clones from a type-1 diabetes patient respond to insulin secretory granule proteins. *Nature*. 1990;345(6276):632-634. doi:10.1038/345632a0

313. Wan X, Zinselmeyer BH, Zakharov PN, et al. Pancreatic islets communicate with lymphoid tissues via exocytosis of insulin peptides. *Nature*. 2018;560(7716):107-111. doi:10.1038/s41586-018-0341-6

314. Bauzá-Martinez J, Heck AJR, Wu W. HLA-B and cysteinylated ligands distinguish the antigen presentation landscape of extracellular vesicles. *Commun Biol*. 2021;4(1):1-11. doi:10.1038/s42003-021-02364-y

315. Magnuson AM, Thurber GM, Kohler RH, Weissleder R, Mathis D, Benoist C. Population dynamics of islet-infiltrating cells in autoimmune diabetes. *Proc Natl Acad Sci USA*. 2015;112(5):1511-1516. doi:10.1073/pnas.1423769112

316. Tan S, Li Y, Xia J, et al. Type 1 diabetes induction in humanized mice. *Proc Natl Acad Sci USA*. 2017;114(41):10954-10959. doi:10.1073/pnas.1710415114

317. Colli ML, Hill JLE, Marroquí L, et al. PDL1 is expressed in the islets of people with type 1 diabetes and is up-regulated by interferons-α and-γ via IRF1 induction. *EBioMedicine*. 2018;36:367-375. doi:10.1016/j.ebiom.2018.09.040

318. Osum KC, Burrack AL, Martinov T, et al. Interferon-gamma drives programmed death-ligand 1 expression on islet β cells to limit T cell function during autoimmune diabetes. *Sci Rep*. 2018;8. doi:10.1038/s41598-018-26471-9

319. Rondas D, D'Hertog W, Overbergh L, Mathieu C. Glucagon-like peptide-1: modulator of β-cell dysfunction and death. *Diabetes Obes Metab*. 2013;15(s3):185-192. doi:10.1111/dom.12165

320. Kielgast U, Asmar M, Madsbad S, Holst JJ. Effect of Glucagon-Like Peptide-1 on α- and β-Cell Function in C-Peptide-Negative Type 1 Diabetic Patients. *J Clin Endocr Metab*. 2010;95(5):2492-2496. doi:10.1210/jc.2009-2440

321. Li Y, Hansotia T, Yusta B, Ris F, Halban PA, Drucker DJ. Glucagon-like peptide-1 receptor signaling modulates beta cell apoptosis. *J Biol Chem*. 2003;278(1):471-478. doi:10.1074/jbc.M209423200

322. Yusta B, Baggio LL, Estall JL, et al. GLP-1 receptor activation improves beta cell function and survival following induction of endoplasmic reticulum stress. *Cell Metab*. 2006;4(5):391-406. doi:10.1016/j.cmet.2006.10.001

323. Kielgast U, Holst JJ, Madsbad S. Antidiabetic actions of endogenous and exogenous GLP-1 in type 1 diabetic patients with and without residual β-cell function. *Diabetes*. 2011;60(5):1599-1607. doi:10.2337/db10-1790

324. von Herrath M, Bain SC, Bode B, et al. Anti-interleukin-21 antibody and liraglutide for the preservation of β-cell function in adults with recent-onset type 1 diabetes: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Diabetes Endocrinol*. 2021;9(4):212-224. doi:10.1016/S2213-8587(21)00019-X

325. Trivedi PM, Graham KL, Scott NA, et al. Repurposed JAK1/JAK2 Inhibitor Reverses Established Autoimmune Insulitis in NOD Mice. *Diabetes*. 2017;66(6):1650-1660. doi:10.2337/db16-1250

326. Gitelman SE, Bundy BN, Ferrannini E, et al. Imatinib therapy for patients with recent-onset type 1 diabetes: a multicentre, randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Diabetes Endocrinol*. 2021;9(8):502-514. doi:10.1016/S2213-8587(21)00139-X

327. Ovalle F, Grimes T, Xu G, et al. Verapamil and beta cell function in adults with recent-onset type 1 diabetes. *Nat Med*. 2018;24(8):1108-1112. doi:10.1038/s41591-018-0089-4

328. Hall JC, Rosen A. Type I interferons: crucial participants in disease amplification in autoimmunity. *Nat Rev Rheumatol*. 2010;6(1):40-49. doi:10.1038/nrrheum.2009.237

329. Ivashkiv LB, Donlin LT. Regulation of type I interferon responses. *Nat Rev Immunol*. 2014;14(1):36-49. doi:10.1038/nri3581

330. Taniguchi T, Ogasawara K, Takaoka A, Tanaka N. IRF family of transcription factors as regulators of host defense. *Annu Rev Immunol*. 2001;19:623-655. doi:10.1146/annurev.immunol.19.1.623

331. Izaguirre A, Barnes BJ, Amrute S, et al. Comparative analysis of IRF and IFN-alpha expression in human plasmacytoid and monocyte-derived dendritic cells. *J Leukoc Biol*. 2003;74(6):1125-1138. doi:10.1189/jlb.0603255

332. Bon AL, Schiavoni G, D'Agostino G, Gresser I, Belardelli F, Tough DF. Type I Interferons Potently Enhance Humoral Immunity and Can Promote Isotype Switching by Stimulating Dendritic Cells In Vivo. *Immunity*. 2001;14(4):461-470. doi:10.1016/S1074-7613(01)00126-1

333. Jego G, Palucka AK, Blanck JP, Chalouni C, Pascual V, Banchereau J. Plasmacytoid Dendritic Cells Induce Plasma Cell Differentiation through Type I Interferon and Interleukin 6. *Immunity*. 2003;19(2):225-234. doi:10.1016/S1074-7613(03)00208-5

334. Luft T, Pang KC, Thomas E, et al. Type I IFNs Enhance the Terminal Differentiation of Dendritic Cells. *J Immunol*. 1998;161(4):1947-1953.

335. Gallucci S, Lolkema M, Matzinger P. Natural adjuvants: Endogenous activators of dendritic cells. *Nat Med*. 1999;5(11):1249-1255. doi:10.1038/15200

336. Marrack P, Kappler J, Mitchell T. Type I interferons keep activated T cells alive. *J Exp Med*. 1999;189(3):521-530. doi:10.1084/jem.189.3.521

337. Biron CA, Nguyen KB, Pien GC, Cousens LP, Salazar-Mather TP. Natural killer cells in antiviral defense: function and regulation by innate cytokines. *Annu Rev Immunol*. 1999;17:189-220. doi:10.1146/annurev.immunol.17.1.189

338. Kolumam GA, Thomas S, Thompson LJ, Sprent J, Murali-Krishna K. Type I interferons act directly on CD8 T cells to allow clonal expansion and memory formation in response to viral infection. *J Exp Med*. 2005;202(5):637-650. doi:10.1084/jem.20050821

339. Curtsinger JM, Valenzuela JO, Agarwal P, Lins D, Mescher MF. Type I IFNs provide a third signal to CD8 T cells to stimulate clonal expansion and differentiation. *J Immunol*. 2005;174(8):4465-4469. doi:10.4049/jimmunol.174.8.4465

340. Qaisar N, Jurczyk A, Wang JP. Potential role of type I interferon in the pathogenic process leading to type 1 diabetes. *Curr Opin Endocrinol Diabetes Obes*. 2018;25(2):94-100. doi:10.1097/MED.0000000000000399

341. Marroqui L, Perez-Serna AA, Babiloni-Chust I, Dos Santos RS. Chapter One - Type I interferons as key players in pancreatic β-cell dysfunction in type 1 diabetes. In: Santin I, Galluzzi L, eds. *International Review of Cell and Molecular Biology*. Vol 359. Pancreatic ß-Cell Biology in Health and Disease. Academic Press; 2021:1-80. doi:10.1016/bs.ircmb.2021.02.011

342. Stefan M, Wei C, Lombardi A, et al. Genetic–epigenetic dysregulation of thymic TSH receptor gene expression triggers thyroid autoimmunity. *Proc Natl Acad Sci USA*. 2014;111(34):12562-12567. doi:10.1073/pnas.1408821111

343. Hasham A, Zhang W, Lotay V, et al. Genetic analysis of interferon induced thyroiditis (IIT): evidence for a key role for MHC and apoptosis related genes and pathways. *J Autoimmun*. 2013;44:61-70. doi:10.1016/j.jaut.2013.04.002

344. Niewold TB, Swedler WI. Systemic lupus erythematosus arising during interferon-alpha therapy for cryoglobulinemic vasculitis associated with hepatitis C. *Clin Rheumatol*. 2005;24(2):178-181. doi:10.1007/s10067-004-1024-2

345. Cacopardo B, Benanti F, Pinzone MR, Nunnari G. Rheumatoid arthritis following PEG-interferon-alfa-2a plus ribavirin treatment for chronic hepatitis C: a case report and review of the literature. *BMC Res Notes*. 2013;6:437. doi:10.1186/1756-0500-6-437

346. Blanco P, Palucka AK, Gill M, Pascual V, Banchereau J. Induction of Dendritic Cell Differentiation by IFN-α in Systemic Lupus Erythematosus. *Science*. 2001;294(5546):1540-1543. doi:10.1126/science.1064890

347. Chehadeh W, Weill J, Vantyghem MC, et al. Increased Level of Interferon-α in Blood of Patients with Insulin-Dependent Diabetes Mellitus: Relationship with Coxsackievirus B Infection. *J Infect D*. 2000;181(6):1929-1939. doi:10.1086/315516

348. Lundberg M, Krogvold L, Kuric E, Dahl-Jørgensen K, Skog O. Expression of Interferon-Stimulated Genes in Insulitic Pancreatic Islets of Patients Recently Diagnosed With Type 1 Diabetes. *Diabetes*. 2016;65(10):3104-3110. doi:10.2337/db16-0616

349. Foulis AK, Farquharson MA, Meager A. Immunoreactive alpha-interferon in insulin-secreting beta cells in type 1 diabetes mellitus. *Lancet*. 1987;2(8573):1423-1427. doi:10.1016/s0140-6736(87)91128-7

350. Nekoua MP, Alidjinou EK, Hober D. Persistent coxsackievirus B infection and pathogenesis of type 1 diabetes mellitus. *Nat Rev Endocrinol*. 2022;18:503-516. doi:10.1038/s41574-022-00688-1

351. Javitt A, Barnea E, Kramer MP, et al. Pro-inflammatory Cytokines Alter the Immunopeptidome Landscape by Modulation of HLA-B Expression. *Front Immunol*. 2019;10:141. doi:10.3389/fimmu.2019.00141

352. Goncalves G, Mullan KA, Duscharla D, et al. IFNγ Modulates the Immunopeptidome of Triple Negative Breast Cancer Cells by Enhancing and Diversifying Antigen Processing and Presentation. *Front Immunol*. 2021;12(645770):645770.

353. McGinty JW, Chow IT, Greenbaum C, Odegard J, Kwok WW, James EA. Recognition of Posttranslationally Modified GAD65 Epitopes in Subjects With Type 1 Diabetes. *Diabetes*. 2014;63(9):3033-3040. doi:10.2337/db13-1952

354. Lundh M, Bugliani M, Dahlby T, et al. The immunoproteasome is induced by cytokines and regulates apoptosis in human islets. *J Endocrinol*. 2017;233(3):369-379. doi:10.1530/JOE-17-0110

355. Kracht MJL, van Lummel M, Nikolic T, et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med*. 2017;23(4):501-507. doi:10.1038/nm.4289

356. Da Silva Xavier G. The Cells of the Islets of Langerhans. *J Clin Med*. 2018;7(3):54. doi:10.3390/jcm7030054

357. James EA, Pietropaolo M, Mamula MJ. Immune Recognition of β-Cells: Neoepitopes as Key Players in the Loss of Tolerance. *Diabetes*. 2018;67(6):1035-1042. doi:10.2337/dbi17-0030

358. Buitinga M, Callebaut A, Sodré FMC, et al. Inflammation-Induced Citrullinated Glucose-Regulated Protein 78 Elicits Immune Responses in Human Type 1 Diabetes. *Diabetes*. 2018;67(11):2337-2348. doi:10.2337/db18-0295

359. Sodré FMC, Bissenova S, Bruggeman Y, et al. Peptidylarginine Deiminase Inhibition Prevents Diabetes Development in NOD Mice. *Diabetes*. 2021;70(2):516-528. doi:10.2337/db20-0421

360. Wang X, Swensen AC, Zhang T, et al. Accurate Identification of Deamidation and Citrullination from Global Shotgun Proteomics Data Using a Dual-Search Delta Score Strategy. *J Proteome Res*. 2020;19(4):1863-1872. doi:10.1021/acs.jproteome.9b00766

361. Alanio C, Lemaitre F, Law HKW, Hasan M, Albert ML. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. *Blood*. 2010;115(18):3718-3725. doi:10.1182/blood-2009-10-251124

362. Shimizu F, Shimizu M, Kamiyama K. Inhibitory Effect of Interferon on the Production of Insulin. *Endocrinology*. 1985;117(5):2081-2084. doi:10.1210/endo-117-5-2081

363. Nordmann TM, Dror E, Schulze F, et al. The Role of Inflammation in β-cell Dedifferentiation. *Sci Rep*. 2017;7(1):6285. doi:10.1038/s41598-017-06731-w

364. Lee H, Lee YS, Harenda Q, et al. Beta Cell Dedifferentiation Induced by IRE1α Deletion Prevents Type 1 Diabetes. *Cell Metab*. 2020;31(4):822-836.e5. doi:10.1016/j.cmet.2020.03.002

365. Jiang H, Canfield SM, Gallagher MP, et al. HLA-E-restricted regulatory CD8(+) T cells are involved in development and control of human autoimmune type 1 diabetes. *J Clin Invest*. 2010;120(10):3641-3650. doi:10.1172/JCI43522

366. Walters LC, Harlos K, Brackenridge S, et al. Pathogen-derived HLA-E bound epitopes reveal broad primary anchor pocket tolerability and conformationally malleable peptide binding. *Nat Commun*. 2018;9(1):3137. doi:10.1038/s41467-018-05459-z

367. Chong C, Marino F, Pak H, et al. High-throughput and Sensitive Immunopeptidomics Platform Reveals Profound Interferonγ-Mediated Remodeling of the Human Leukocyte Antigen (HLA) Ligandome. *Mol Cell Proteomics*. 2018;17(3):533-548. doi:10.1074/mcp.TIR117.000383

368. Girdlestone J, Isamat M, Gewert D, Milstein C. Transcriptional regulation of HLA-A and -B: differential binding of members of the Rel and IRF families of transcription factors. *Proc Natl Acad Sci USA*. 1993;90(24):11568-11572. doi:10.1073/pnas.90.24.11568

369. Anderson AM, Landry LG, Alkanani AA, et al. Human islet T cells are highly reactive to preproinsulin in type 1 diabetes. *Proc Natl Acad Sci USA*. 2021;118(41):e2107208118. doi:10.1073/pnas.2107208118

370. Thomaidou S, Slieker RC, van der Slik AR, et al. Long RNA Sequencing and Ribosome Profiling of Inflamed β-Cells Reveal an Extensive Translatome Landscape. *Diabetes*. 2021;70(10):2299-2312. doi:10.2337/db20-1122

371. Mann SE, Zhou Z, Landry LG, et al. Multiplex T Cell Stimulation Assay Utilizing a T Cell Activation Reporter-Based Detection System. *Front Immunol*. 2020;11. doi:10.3389/fimmu.2020.00633

372. Kim SK, Woodcroft KJ, Khodadadeh SS, Novak RF. Insulin Signaling Regulates γ-Glutamylcysteine Ligase Catalytic Subunit Expression in Primary Cultured Rat Hepatocytes. *J Pharmacol Exp Ther*. 2004;311(1):99-108. doi:10.1124/jpet.104.070375

373. Domínguez C, Ruiz E, Gussinye M, Carrascosa A. Oxidative Stress at Onset and in Early Stages of Type 1 Diabetes in Children and Adolescents. *Diabetes Care*. 1998;21(10):1736-1742. doi:10.2337/diacare.21.10.1736

374. Darmaun D, Smith SD, Sweeten S, Sager BK, Welch S, Mauras N. Evidence for Accelerated Rates of Glutathione Utilization and Glutathione Depletion in Adolescents With Poorly Controlled Type 1 Diabetes. *Diabetes*. 2005;54(1):190-196. doi:10.2337/diabetes.54.1.190

375. Kohnert KD, Jahr H, Schmidt S, Hahn HJ, Zühlke H. Demonstration of insulin degradation by thiol-protein disulfide oxidoreductase (glutathione-insulin transhydrogenase) and proteinases of pancreatic islets. *Biochimica et Biophysica Acta (BBA) - Enzymology*. 1976;422(2):254-259. doi:10.1016/0005-2744(76)90136-4

376. Trujillo JA, Croft NP, Dudek NL, et al. The Cellular Redox Environment Alters Antigen Presentation. *J Biol Chem*. 2014;289(40):27979-27991. doi:10.1074/jbc.M114.573402

377. Chen W, Yewdell JW, Levine RL, Bennink JR. Modification of Cysteine Residues In Vitro and In Vivo Affects the Immunogenicity and Antigenicity of Major Histocompatibility Complex Class I–restricted Viral Determinants. *J Exp Med*. 1999;189(11):1757-1764. doi:10.1084/jem.189.11.1757

378. Pierce RA, Field ED, Haan JMM den, et al. Cutting Edge: The HLA-A*0101-Restricted HY Minor Histocompatibility Antigen Originates from DFFRY and Contains a Cysteinylated Cysteine Residue as Identified by a Novel Mass Spectrometric Technique. *J Immunol*. 1999;163(12):6360-6364.

379. Kronenberg-Versteeg D, Eichmann M, Russell MA, et al. Molecular Pathways for Immune Recognition of Preproinsulin Signal Peptide in Type 1 Diabetes. *Diabetes*. 2018;67(4):687-696. doi:10.2337/db17-0021

380. Hebbandi Nanjundappa R, Ronchi F, Wang J, et al. A Gut Microbial Mimic that Hijacks Diabetogenic Autoreactivity to Suppress Colitis. *Cell*. 2017;171(3):655-667.e17. doi:10.1016/j.cell.2017.09.022

381. Tai N, Peng J, Liu F, et al. Microbial antigen mimics activate diabetogenic CD8 T cells in NOD mice. *J Exp Med*. 2016;213(10):2129-2146. doi:10.1084/jem.20160526

382. Lanz TV, Brewer RC, Ho PP, et al. Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature*. 2022;603(7900):321-327. doi:10.1038/s41586-022-04432-7

383. Bentzen AK, Marquard AM, Lyngaa R, et al. Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat Biotechnol*. 2016;34(10):1037-1045. doi:10.1038/nbt.3662

384. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14(4):417-419. doi:10.1038/nmeth.4197

385. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22(9):1760-1774. doi:10.1101/gr.135350.111

386. Segerstolpe Å, Palasantza A, Eliasson P, et al. Single-Cell Transcriptome Profiling of Human Pancreatic Islets in Health and Type 2 Diabetes. *Cell Metab*. 2016;24(4):593-607. doi:10.1016/j.cmet.2016.08.020

387. Nyalwidhe JO, Gallagher GR, Glenn LM, et al. Coxsackievirus-Induced Proteomic Alterations in Primary Human Islets Provide Insights for the Etiology of Diabetes. *J Endocr Soc*. 2017;1(10):1272-1286. doi:10.1210/js.2017-00278

388. Nyalwidhe JO, Jurczyk A, Satish B, et al. Proteomic and Transcriptional Profiles of Human Stem Cell-Derived β Cells Following Enteroviral Challenge. *Microorganisms*. 2020;8(2). doi:10.3390/microorganisms8020295

389. Gebhard JR, Perry CM, Harkins S, et al. Coxsackievirus B3-induced myocarditis: perforin exacerbates disease, but plays no detectable role in virus clearance. *Am J Pathol*. 1998;153(2):417-428. doi:10.1016/S0002-9440(10)65585-X

390. Ramsingh AI, Lee WT, Collins DN, Armstrong LE. T Cells Contribute to Disease Severity during Coxsackievirus B4 Infection. *J Virol*. 1999;73(4):3080-3086. doi:10.1128/JVI.73.4.3080-3086.1999

391. Stetson DB, Medzhitov R. Type I Interferons in Host Defense. *Immunity*. 2006;25(3):373-381. doi:10.1016/j.immuni.2006.08.007

392. Cao K, Hollenbach J, Shi X, Shi W, Chopek M, Fernández-Viña MA. Analysis of the frequencies of HLA-A, B, and C alleles and haplotypes in the five major ethnic groups of the United States reveals high levels of diversity in these loci and contrasting distribution patterns in these populations. *Human Immunology*. 2001;62(9):1009-1030. doi:10.1016/S0198-8859(01)00298-1

393. Verkerk T, Koomen SJI, Fuchs KJ, Griffioen M, Spaapen RM. An unexplored angle: T cell antigen discoveries reveal a marginal contribution of proteasome splicing to the immunogenic MHC class I antigen pool. *Proc Natl Acad Sci USA*. 2022;119(29):e2119736119. doi:10.1073/pnas.2119736119

394. Park JE, Miller Z, Jun Y, Lee W, Kim KB. Next-generation proteasome inhibitors for cancer therapy. *Transl Res*. 2018;198:1-16. doi:10.1016/j.trsl.2018.03.002

395. Kirk CJ, Muchamuel T, Wang J, Fan RA. Discovery and Early Clinical Development of Selective Immunoproteasome Inhibitors. *Cells*. 2021;11(1):9. doi:10.3390/cells11010009

396. Łącki MK, Startek MP, Brehmer S, Distler U, Tenzer S. OpenTIMS, TimsPy, and TimsR: Open and Easy Access to timsTOF Raw Data. *J Proteome Res*. 2021;20(4):2122-2129. doi:10.1021/acs.jproteome.0c00962

397. Kong AT, Leprevost FV, Avtonomov DM, Mellacheruvu D, Nesvizhskii AI. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry–based proteomics. *Nat Methods*. 2017;14(5):513-520. doi:10.1038/nmeth.4256

398. Parker R, Tailor A, Peng X, et al. The Choice of Search Engine Affects Sequencing Depth and HLA Class I Allele-Specific Peptide Repertoires. *Mol Cell Proteomics*. 2021;20:100124. doi:10.1016/j.mcpro.2021.100124

399. Brennick CA, George MM, Srivastava PK, Karandikar SH. Prediction of cancer neoepitopes needs new rules. *Semin Immunol*. 2020;47:101387. doi:10.1016/j.smim.2020.101387

400. Ebrahimi-Nik H, Moussa M, Englander RP, et al. Reversion analysis reveals the in vivo immunogenicity of a poorly MHC I-binding cancer neoepitope. *Nat Commun*. 2021;12(1):6423. doi:10.1038/s41467-021-26646-5

401. Nakayasu ES, Syed F, Tersey SA, et al. Comprehensive Proteomics Analysis of Stressed Human Islets Identifies GDF15 as a Target for Type 1 Diabetes Intervention. *Cell Metab*. 2020;31(2):363-374.e6. doi:10.1016/j.cmet.2019.12.005